PSY
UNIVERSITY OF MICHIGAN
JUN 19 1951
SCIENCE LIBRARY

Vol. 48, No. 3 May, 1951

Psychological Bulletin

CONTENTS

GENERAL REVIEWS:

Salivary pH as a Psychophysiological Variable: J. C. GILCHRIST AND ERNEST FURCHTGOTT, 193.

Research with the Wechsler-Bellevue Test: 1945–1950: ALBERT I. RABIN AND WILSON H. GUERTIN, 211.

NOTES:

A Psychophysical Method Employing a Modification of the Müller-Urban Weights: STEPHEN HARRISON AND MARGARET J. HARRISON, 249.

Some Comments on the Controlling of Differences in Responses on the Rorschach Test: ROBERT M. EICHLER, 257.

BOOK REVIEWS:

BORING'S *A history of experimental psychology:* C. H. GRAHAM, 260.

HOLE'S *Psychology: Its principles and applications:* FREDERICK A. COURTS, 263.

GUTHRIE AND POWERS' *Educational psychology:* FRANK K. SHUTTLEWORTH, 264.

ANASTASI AND FOLEY'S *Differential psychology:* DALE B. HARRIS, 265.

RUPPOCK'S *Group guidance: Principles, techniques, and evaluation:* WILBUR S. GREGORY, 267.

PLANT'S *The envelope: A study of the impact of the world upon the child:* RICHARD A. LITTMAN, 268.

EDWARDS' *Experimental design in psychological research:* IRVIN L. CHILD, 269.

OCHRAN AND COX'S *Experimental designs:* DAVID A. GRANT, 272.

ERRIS' *Some relations between vision and audition:* FRANK A. GELDARD, 273.

STORE'S *The nature-nurture controversy:* JOHN E. ANDERSON, 273.

RLOCK'S *Child development:* S. O. ROBERTS, 275.

N BONIN'S *Essay on the cerebral cortex:* HENRY W. NISSEN, 276.

TEN'S *Surveys, polls and samples: practical procedures:* ELI S. MARKS, 277.

RMAN'S *Theory and practice of psychological testing:* FREDERICK B. DAVIS, 279.

PNER'S *Psychology applied to life and work:* IRVING R. WESCHLER, 280.

ROMM-REICHMANN'S *Principles of intensive psychotherapy:* O. H. MOWRER, 281.

EEVER, TURRELL, AND WEITZEL'S *Principles and techniques of guidance:* JOSEPH V. HANNA, 283.

'S *Mental tests in clinics for children:* HELMER R. MYKLEBUST, 284.

HAIN AND MACLEAN'S *General clinical counseling:* RALPH F. BERDIE, 285.

NDIS AND BOLLES' *Textbook of abnormal psychology:* ROBERT P. FISCHER, 286.

TEINS' *Hypnotherapy of war neuroses:* RALPH D. NORMAN, 287.

PUBLISHED BI-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

LYLE H. LANIER, Editor
University of Illinois

MARGARET K. HARLOW, Managing Editor

CONSULTING EDITORS

STUART H. BRITT
McCann-Erickson, Inc., New York
DORWIN CARTWRIGHT
University of Michigan
FRANK A. GELDARD
University of Virginia
JAMES J. GIBSON
Cornell University
DAVID A. GRANT
University of Wisconsin
WILLIAM T. HERON
University of Minnesota
ERNEST R. HILGARD
Stanford University
WILLIAM A. HUNT
Northwestern University
JEAN WALKER MACFARLANE
University of California
DONALD G. MARQUIS
University of Michigan
JOHN T. METCALF
University of Vermont
JAMES G. MILLER
University of Chicago
NEAL E. MILLER
Yale University
HELEN PEAK
University of Michigan
ROBERT R. SEARS
Harvard University
ROBERT L. THORNDIKE
Teachers College, Columbia University

The Psychological Bulletin contains evaluative reviews of the literature in various fields of psychology, methodological articles, critical notes, and book reviews. *This* JOURNAL *does not publish reports of original research or original theoretical articles.*

Editorial communications, manuscripts and book reviews should be sent to Lyle H. Lanier, Department of Psychology, University of Illinois, Urbana, Illinois.

Preparation of articles for publication. Authors are strongly advised to follow the general directions given in the article by Anderson and Valentine, "The preparation of articles for publication in the journals of the American Psychological Association" (*Psychological Bulletin*, 1944, 41, 345–376). Special attention should be given to the section on the preparation of the bibliography (pp. 363–372), since this is a particular source of difficulty in long reviews of research literature. *All copy must be double-spaced, including the bibliography.*

Reprints. Fifty reprints are given, if requested, to contributors of articles, notes and special reviews. Five copies of the JOURNAL are supplied gratis to the authors of book reviews.

Business communications—including subscriptions, orders of back issues and changes of address—should be sent to the American Psychological Association, 1515 Massachusetts Avenue, N. W., Washington 5, D. C.

Annual subscription $7.00 (Foreign $7.50). Single copies, $1.25.

PUBLISHED BI-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1515 Massachusetts Ave., N.W., Washington 5, D.C.

Entered as second class mail matter at the post office at Washington, D.C., under the act of March 3, 1879. Additional entry at the post office at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in Section 538, act of February 28, 1925, authorized August 6, 1947.

Copyright, 1951, by The American Psychological Association, Inc.

Vol. 48, No. 3 May, 1951

Psychological Bulletin

SALIVARY pH AS A PSYCHOPHYSIOLOGICAL VARIABLE[1]

J. C. GILCHRIST[2] AND ERNEST FURCHTGOTT[3]
University of California, Los Angeles

There are many suggestions in the literature of the biological sciences that the pH of saliva[4] may be a valuable tool for the diagnosis and prediction of behavior. These suggestions arise from two major considerations. First, the salivary glands are dually innervated by the sympathetic and the parasympathetic nervous systems and, consequently, may furnish indices of autonomic balance. Secondly, it is well known that extreme shifts in the pH of blood bring about behavioral changes ranging from tetany in alkalinosis to coma in acidosis. The blood, however, is well buffered, and minor shifts in the organism's acid-base balance do not appear in blood pH determination. Since saliva is less well buffered, it seems possible that its pH may serve as an index of the acid-base balance of the fluid matrix. Wenger's (58) recent study of Air Forces personnel supports at least the first of these two possibilities. That study showed salivary pH to be significantly related to other autonomic nervous system functions and demonstrated that it can be used to estimate autonomic balance. Such findings indicate that a review of the literature pertaining to salivary pH may be valuable.

Starr (49), in his paper in 1922, presented the results of earlier investigators dating back to Donné's observations in 1833. The present review attempts to cover the work done in the field since 1922. Particular attention is devoted to recent reports which, since the advent of electrometric methods, provide the most reliable data.

[1] The authors wish to express their appreciation to Dr. M. A. Wenger for suggesting this review and assisting them in its preparation.

[2] Now at the University of Wisconsin.

[3] Now at the University of Tennessee.

[4] pH refers to the relative concentrations of H^+ and OH^- ions within a substance and thereby its actual acidity or alkalinity. The pH value is the negative logarithm (the sign is dropped for convenience) of the concentration of the hydrogen ion in gram atoms per liter. A pH of less than 7 denotes acidity, whereas a pH of greater than 7 denotes alkalinity.

Methods of Collection and Measurement

It should be pointed out at once that a majority of reports dealing with determinations of salivary pH have made use of data obtained from mixed whole saliva. A few investigators, however, have collected the secretions of individual glands. In most cases they worked with the parotids and employed a Lashley disk (**34**) attached by negative pressure over the orifice of the gland. There are several modifications of this disk, such as an added third baffle, used by Moore and Marcuse (**38**), but it is doubted that the use of this or other variations perceptibly increases the validity of the method. In a recent study, Schmidt-Nielson (**45**) used two rubber cups which were attached to the membrane around the papillae of the parotid ducts, while the submaxillary flow was sucked directly from the floor of the mouth just under the tongue. When mixed whole saliva is collected, it is usually done by use of negative pressure devices or by having the subject expectorate.

Saliva from individual or all glands is usually collected under oil, toluene, or some other inert substance to prevent the loss of CO_2, since it was shown by Starr (**50**) that this loss causes a rise in salivary pH. Schmidt-Nielsen (**45**) notes that the loss of CO_2 from saliva is the greatest in the first 90 seconds and that toluene or paraffin oil are not very effective in preventing this loss inasmuch as they have a CO_2 diffusion rate only four times slower than water. Wenger, Furchtgott, and Voas (**59**) measured electrometrically the salivary pH of 12 subjects, with the saliva collected first under toluene and later measured directly in the mouth. For the direct measurements the mean and standard deviation were 6.45 and 0.36, respectively, and under toluene they were 6.50 and 0.40, respectively. The difference involved is not statistically significant. Schmidt-Nielsen used a syringe pipette into which saliva was sucked directly from a rubber cup and the dead space in the syringe was filled with mercury to prevent CO_2 loss. Most workers have not controlled the initial stages of their collection so carefully, and it will be seen later that they have obtained pH values higher than those obtained by Schmidt-Nielsen. Certain varying interpretations of results become possible if these factors are kept in mind, but lack of rigorous initial control need not invalidate the data from a single study as long as the worker maintains a constant technique. In fact, if it be assumed that the initial CO_2 loss is similar for all samples, it has been shown (**59**) that if pH determinations are made within five minutes after collection, it is unnecessary to use any protecting device for CO_2, unless an accuracy of greater than 0.01 pH units is desired. Gaskill (**22**) even reports that

saliva stored in a small Pyrex bottle fitted with an airtight cap remains unchanged in pH for 48 hours.

As Starr (**51**) and others have pointed out, neither titration nor colorimetric indicators provide entirely satisfactory salivary pH determinations. Titration must be questioned because the concentration of hydrogen ions in a given volume of liquid, rather than the quantity of alkali replaced in the course of titration, determines the significant biological properties of most fluids. The efficiency of the liquid colorimetric indicators is impaired due to the wide variation in the turgidity of saliva. This, of course, makes for varying degrees of chromatic saturation in the indicator, since these methods require the use of a color comparator with the same light source used for illuminating both the standards and the material to be tested.

Glaubiger and Falk (**23**) concluded from a series of investigations using electrometric and colorimetric techniques that the latter may be used if they are checked regularly with standard solutions and if determinations are needed only within 0.3 pH. Paper colorimetric indicators, in the main, are designed to measure only within the nearest 0.5 pH unit and have only a short range. The Wulff paper indicator strip is designed to be used for turbid, viscous materials with an accuracy of 0.1 pH unit, although by the use of standard buffer solutions it is claimed that the accuracy can be extended to 0.05 pH unit. No data have been found which compare this technique with other colorimetric methods.

Most investigators regard electrometric techniques as the only ones providing accurate pH determinations, and the glass electrode is the most frequently used today. Soyenkoff and Hinch (**48**) claim that the difference in using a quinhydrone electrode is less than 0.02 pH, but Wessinger (**56**) notes that the potential drift is greater with increased deviation of the oxidation-reduction potential from zero. He describes a vacuum tube potentiometer which eliminates this drift.

Microtechniques are now being used by some workers, and studies in situ appear possible. Schmidt-Nielsen (**45**) recently published a method for microdetermination which produces fairly reliable results with less than 0.10 ml. of saliva required for determination. A Billman-type (**7**) quinhydrone capillary electrode is employed, filled with about 0.05 ml. of a standard buffer solution. Claff and Swenson (**15**) describe another technique for the determination of samples of biological fluids as small as 0.005 ml. Nims's (**40**) glass electrode and apparatus used to measure the pH of the brain surface in vivo could be adopted to de-

termine the pH of salivary glands in situ. Stephan (**54**) has employed an antimony electrode to measure the pH of the tooth surface, which might also be modified to measure salivary pH. Recent work (**59**) indicates that direct oral measurements are possible with small commercial glass electrodes.

pH Values for Human Saliva

Until recently, most authors have considered saliva as having an alkaline reaction. (See Zoethout and Tuttle [**62**, p. 343] for a typical statement of this position.) Table I shows the results obtained in ten of

TABLE I

Reported Values of Salivary pH

Investigator	*Collection Method*	*Measurement Method*	*Samples and Subjects*	*pH Values Reported*
Brawley (12)	Expectoration	Colorimetric (liquid)	1 sample from each of 1826 males and 1578 females, aged 3 weeks–101 years	M =6.75 (Males =6.76; Females =6.73) Mdn =6.84
Carlson & McKinstry (14)	Expectoration	Colorimetric (liquid)	1 sample from each of 127 male college students	M =6.85 Range =6.66 to 7.02
Eisenbrandt (19)	Expectoration	Electrometric	276 samples from 5 subjects at 9, 11, 1, 3, and 5 o'clock during the day	M =6.72, σ =.242
Eisenbrandt (20)	Expectoration	Electrometric	1552 samples from 7 subjects at hourly intervals during the day for 1 week of each month for 1 year	M =6.64, σ_M =.006
Gans (21)	Expectoration	Colorimetric (liquid)	1 sample from each of 90 subjects	Range =6.0 to 7.0
Reindel (42)		Colorimetric	50 cases with "normal" gastric acidity	39 of 50 cases ranged between 6.9 and 7.1
Rich (43)	Expectoration	Electrometric	3 samples from each of 57 subjects 2 hours after meals	M =6.82 Range =5.88 to 7.38, with 31% of observations >7.0
Starr (50)	Expectoration	Colorimetric (liquid)	610 samples from 228 subjects	M =6.6 Mdn =6.6 Range =5.95 to 7.25, with 86.6% of the samples between 6.55 and 7.00
Wenger (58)	Drawn from oral cavity under negative pressure	Colorimetric (paper)	1 sample from each of 488 aviation students; 225 operational fatigue cases; 98 hospitalized psychoneurotic patients; 166 army returnees without operational fatigue	M =7.2, σ =0.4 M =7.1, σ =0.5 M =6.9, σ =0.7 M =6.9, σ =0.6
Wenger, Furchtgott, & Voas (59)	Drawn from oral cavity under negative pressure	Electrometric	1 sample from each of 69 college students (both sexes)	M =6.45, σ =0.41

the most complete studies of mixed unstimulated[5] saliva in normal human subjects. Although several of the studies are incompletely described statistically, and only eight of the ten report even so much as the mean, seven of these eight studies report fairly consistent results. The obtained means were 6.75, 6.85, 6.72, 6.64, 6.82, 6.78, and 6.45. Three of these studies used colorimetric techniques, while four sets of data were obtained electrometrically. Of the remaining experiments only one reports values consistently on the alkaline side. This group of studies indicates that the pH of mixed unstimulated human saliva is slightly acid, and places its actual mean value at some point between 6.45 and 6.85. The higher values reported by Wenger (**58**) remain to be explained.

Little evidence is available on the comparative contributions of the different glands. The Schmidt-Nielsen study (**46**) provides the only available comparison of parotid and submaxillary pH measured in the same subjects. The average parotid pH for seven subjects, tested four times daily, and measured electrometrically, was 5.81, and that for submaxillary saliva was 6.39. Winsor and Korchin (**60**), also using electrical methods, report 5.83 as the mean parotid pH for several hundred subjects. As we have seen, these are considerably lower values for mixed saliva than those of other investigators when similar measuring techniques are employed. It would seem that the various salivary glands make differing contributions to the whole, a conclusion which is not surprising since the parotids are composed largely of serous-type cells, while the submaxillary and sublingual glands contain mucous-type cells. It remains for future research to determine the differential participation of the different glands on the results obtained for mixed saliva.

Seasonal Variations

In addition to the problem of measurement itself, there are questions concerning the effects of experimental conditions on salivary pH. One of the questions is that of seasonal variations. Eisenbrandt (**20**), measuring 1552 specimens from seven subjects tested at bihourly intervals for one week of each month for one year, found by an analysis of variance that the pH of the group was significantly variable from month to month at each bihourly period.

While seasonal variations were observed in the above study, no attempt was made to correlate them with atmospheric conditions. Indirectly, a study by Berg *et al.* (**3**) may be compared. They found that the pH of venous blood measured in a longitudinal study of six dogs and

[5] "Unstimulated" means here that the salivary flow was not elicited by the injection of stimulants intravenously or by placing a stimulating material in the oral cavity.

six human beings for one year correlated with a significance above the 1 per cent level with barometric pressure and air temperature. The correlations were:

	Barometric Pressure	*Air Temperature*
Human venous blood.....	.77	.47
Canine venous blood.....	.69	.45

The relationship between air temperature and pH was marked only at higher temperatures. Bearing more directly upon the problem is the preliminary study by Wenger (**58**) on 180 Air Forces cadets over a four-month period. He obtained a correlation of +.21 between salivary pH and air temperature, and +.06 between salivary pH and barometric pressure when the relationship between barometric pressure and air temperature was partialed out.

Insofar as there is a positive relationship between salivary output and pH, the effect of seasonal changes on the former is pertinent to the review. An earlier study by Wenger (**57**) on 51 children and again on 23 subjects (**57**) reported greater salivary output in the summer than in the winter; however, the differences were not statistically significant (t values reported were 1.40 and 0.97, respectively). Wainwright and Becks's study (**55**) also furnishes evidence on this point. They investigated rate of flow of saliva in 50 subjects over a period of three years. There were no seasonal changes found. From these studies it can be concluded that there is probably only insignificant seasonal variation in salivary pH except at higher temperatures, where changes can be expected due to the physicochemical relationships between temperature and pH.

Diurnal Variations

Henderson and Millet (**27**) state that saliva is most acid in the morning upon arising, begins to approach neutrality immediately, rises to alkalinity after breakfast, drops to acidity afterwards, rises again after ingestion of the noon meal, following which there is another drop. The determinations were done colorimetrically using the expectorated saliva of 13 subjects who were tested during the course of a whole day, and numerous others who were tested for half-days during a period of 15 days. McClelland (**36**) previously found somewhat the same type of diurnal variation in four subjects tested on four days, although his absolute values differ from those of Henderson and Millet (**27**). Measuring expectorated saliva colorimetrically, he found the lowest pH to occur after arising, followed by a gradual increase reaching its peak just after lunch, which was followed by a decline of the pH values. Smith

(47), who employed a quinhydrone electrode on an unspecified number of subjects, obtained data similar to those of Henderson and Millet (27). The lowest pH was recorded in the morning; this was followed by a rise reaching its maximum about four hours after breakfast and then the pH dropped again.[6] Eisenbrandt (19) examined 276 specimens of five subjects electrometrically and found that the lowest pH occurred in the morning, followed by an increased pH at 11 A.M. and a peak at 5 P.M. Grossman and Brickman (25) studied three groups which were differentiated on the basis of diurnal and nocturnal activity. Group 1, active at night, had a diurnal pH of 6.7 and a nocturnal pH of 6.7. Group 2, which had eight hours of sleep at night, had a diurnal pH of 6.7 and a nocturnal pH of 6.1. Group 3, consisting of ambulant and bed patients, had a diurnal pH of 6.3 and a nocturnal pH of 6.1. The saliva consisted of mixed expectorated specimens, and the pH was determined with a quinhydrone electrode. In contrast to these studies, Schmidt-Nielsen (46) using his technique which has been described previously, examined the parotid saliva of five subjects four times daily and the saliva of two subjects seven times daily at regular intervals during the 24-hour period and found no diurnal or nocturnal variations in the pH of the saliva. However, he did find fluctuations in the pH of submaxillary saliva. Wenger (58), in his study of 180 Air Forces cadets, obtained an r of +.19 between time of day and salivary pH. In a more recent study, Wenger, Furchtgott, and Voas (59) found that with 39 subjects whose salivary pH was measured at 10 A.M. and again at 2 P.M., the difference between the means was insignificant (C.R. = .5) and that the two measurements correlated with an r of +.69. No attempt was made to control the subjects' usual eating habits, but subjects were instructed to refrain from eating for two hours prior to measurement, and presumably most of the measurements were obtained after the ingestion of the morning and the noon meals, respectively.

Although these results are not entirely consistent, it would seem that when diurnal variations in salivary pH are found, the values are not specifically related to the time of day but, rather, to physical activity and to the ingestion of food, both of which may increase pH temporarily.

DIET

Diet apparently has no effect upon salivary pH. Smith (47) fed an

[6] Lashley (34) found that the presence of food in the stomach increases salivary flow. Coupled with the relationship between rate of flow and pH (see pp. 201–202) this finding would account for the peaks in the above studies.

unspecified number of subjects a high protein diet, a diet high in $NaHCO_3$ and sodium citrate, and a high carbohydrate diet. The salivary pH of the subjects was independent of the diet fed, in contrast to the urinary pH, which was affected by the diets. Clark *et al.* (16) found that the acid-base balance of the diet bore little or no relationship to the quantity of acid or base found in the blood or in unstimulated saliva. Schmidt-Nielsen (46), measuring salivary pH soon after the ingestion of a meal, found that it was independent of the character of the meal.

Chemical Stimulants in the Oral Cavity

The results of placing various chemical substances in the oral cavity are consistent in their effects upon salivary pH. The inverse relationship of salivary pH to the pH of the stimulating substance has been demonstrated in animal and human subjects by Popow and Kudrjawzew (41), Bienka and Sczepansky (5), Bienka (6), Krassusky *et al.* (33), and Winsor and Korchin (60). These latter authors, however, state that the salivary pH remained directly proportional to the rate of flow of saliva, and thus the pH of the stimulating substance may have only an indirect effect on salivary pH by controlling salivary flow which, as will be described later, bears a direct relationship to salivary pH.

Jaw Movements

Many investigators have noted that jaw movements increase salivary pH. McClelland (36), who used colorimetric measures of expectorated saliva, found an increase of pH after movements of the jaw. Sneezing, coughing, and talking produce a similar rise. Schmidt-Nielsen (46) recorded an increase from 5.62 to 6.68 pH units after coughing. Carlson and McKinstry (14) also found an increased pH after jaw movements, while Winsor and Korchin (60), working with parotid glands, found that the jaw movement, regardless of the nature of the chewed object, produced an increased rate of flow which was almost directly proportional to pH.

Wenger, Furchtgott, and Voas (59) measured electrometrically the pH of saliva in the oral cavity for 69 subjects and compared the mean value obtained by direct measurement with that obtained by a vacuum suction collection method in which chewing movements were involved. The mean of the former was 6.36, and that of the latter was 6.45. Although the difference is not statistically significant, it is in the expected direction.

Rate of Salivary Flow

As was indicated above Winsor and Korchin (60), working with human subjects aged 13–23, found a direct relationship between rate of flow and salivary pH. Determining the pH electrometrically, they found an increased rate of secretion of the parotid gland is paralleled by an increased salivary pH. Conversely, a decreased rate of parotid secretion also produces a more acid saliva. Wenger (58) also obtained a positive correlation (r of +.44) between salivary pH and output in his study of 488 aviation students, and for 201 *Operational Fatigue* patients, an r of +.30 was obtained. All pH determinations in this study were by means of nitrazene paper indicators.

These results agree with the observations that, in acidosis (4) and fatigue (1), when the blood lactate is high and there is an interchange between this substance and the gland bicarbonate, the rate of secretion is low and the pH is also decreased.

The above results are at variance, however, with those obtained on subhuman subjects. Brassfield (8) found that the secretory rate in dogs is inversely proportional to pH. Increasing the cellular acidity of the gland inhibits the destruction of acetylcholine, resulting in an increased secretion. The effect is similar to observations in other cholinergic systems, according to Brassfield, who further states that "variations in base with variations in rate of secretion, the type of metabolism within the gland, are all important factors in determining the reactions of saliva" (p. 48). The work of Bienka (5), Bienka and Sczepansky (6), and Bronk and Gesell (13) agrees with that of Brassfield—i.e., that pH is inversely proportional to rate of flow. Kesztyüs and Martin (30) also report that increased secretion is accompanied by a decreased content of Na and therefore of total base, which means that the pH will be lowered. The above studies were conducted primarily on the submaxillary glands of dogs.

Two possible reasons present themselves for the discrepancies between the human studies and those on other animals. The first is that the studies of nonhuman animals elicited saliva by artificial means (e.g., drugs, electrical current, etc.) which may produce different results. In fact, Baxter (2) has demonstrated that the composition of saliva after pilocarpine stimulation is different from that obtained by the use of an electrical current.

Furthermore, the studies on human subjects employed either "mixed saliva" or secretions of the parotids only, while the nonhuman data were obtained from the submaxillary glands. The differences, therefore, may be due to the differential activity of these glands upon

stimulation. Stavraky (53), for example, has shown that the two glands react differently to adrenaline following the previous administration of pilocarpine; and it is even more appropriate to note the results of Holzlöhner and Airapetianz (38), who found that sympathetic stimulation of the submaxillary glands in dogs produced an increase in secretion. This is in contrast to the parotids, where sympathetic stimulation results in an inhibition of salivary flow. The above findings, taken in conjunction with the work of Kesztyüs and Martin (31) (who report a lower pH for sympathetically than for parasympathetically stimulated submaxillary saliva in dogs), give a plausible explanation for the apparent contradictions between the human and other animal studies as regards the relationship between rate of flow and pH. The direct relationship between rate of flow and pH observed in human subjects is probably due to an increased pH with parasympathetic stimulation of the parotids, which also produces an increased rate of flow. The inverse relationship between rate of flow and pH in dogs is due to sympathetic stimulation of the submaxillary glands, which raises the rate of flow while at the same time it lowers the pH. From the available data it would appear, therefore, that the relationship between rate of flow and salivary pH is an indirect one, expressing mainly the increased pH resulting from parasympathetic stimulation, which in the case of the parotids or mixed saliva is also associated with increased rate of flow.

pH of the Blood and Other Body Fluids

Hugh and Marenczi (29), Gregersen and Ingalls (24), and DeBeer and Wilson (18) found a rise in salivary pH in animals following a slow or moderate intravenous injection of hydrochloric acid. However, according to Benetate and Oprisiu (4), when the concentration of acid is increased, or in acidosis, or after exercise, the pH of saliva follows that of blood, i.e., it drops. Leung (35) also reported that the pH of human saliva decreases under a lowered oxygen tension, and similar results were previously reported by Starr (50).

Brassfield and Hong (11) found no change in either flow or pH after intravenous injections of lactic and hydrochloric acids, while sodium carbonate and sodium bicarbonate decreased the flow and increased the pH of dogs' submaxillary saliva, and ammonium hydroxide and ammonium carbonate increased the flow and decreased pH, according to Krassusky *et al.* (33) and Brassfield *et al.* (10). The latter authors relate the apparently contradictory results obtained with the ammonium compounds to their anti-cholinesterase activity. Previously, Brassfield and Hands (9) reported the same conclusions. They also noted that

procedures which caused a greater pH increase in blood than in saliva also decreased the salivary flow, while in those cases in which there was a sharper rise in the salivary than in the blood pH, the rate of salivary flow was increased. Overventilation, for example, produced an increase in both blood and salivary pH, with a lesser increase in the former and an accompanying increased salivary flow. Again, in 1945, Brassfield (8) described some experiments in which he stimulated the submaxillary glands of dogs continuously with pilocarpine and confirmed his earlier findings. He says of this work: "The lowering of the CO_2 tension of the blood increases the CO_2 gradient between the gland cell and the blood, thereby providing a means for rapidly decreasing the CO_2 tension of saliva and the increasing salivary pH. Since the change in salivary pH is much smaller than the pH change in the arterial blood, it is obvious that some other factor is also involved in determining the salivary pH" (p. 47). The other factors, according to Brassfield, are the rate of secretion and the type of metabolism in the gland.

Although this evidence may appear to vitiate the hypothesis that salivary pH does not follow blood pH, actually it does not do so, for the typical finding has been that salivary pH and blood pH do not parallel one another until drastic changes have been made in the blood pH. Apparently, a restatement of the position is necessary. Salivary pH will follow blood pH only when extensive or long-sustained changes occur in the blood. With lesser or phasic changes, salivary pH may even shift in the opposite direction, depending upon other effects of the particular substance eliciting the change.

Fatigue

Since Starr (51) reported a lowering of pH in subjects whose energy level was seemingly lowered, a number of experimenters have described similar findings. Henderson and Millet (27) recorded a drop of pH in one fatigued subject; and Baker and Eye (1), measuring pH electrometrically in four subjects, also noted a drop during rest after exercise. Winsor and Korchin (60) tested several subjects electrometrically and found a drop in pH after exercise, and Nekrasov and Khranilova (39) also reported similar results. These data are in accord with the experiments on HCl injections and the expectations with regard to acidosis.

Other Physiological Measures

Korchin and Winsor (32) tested the rate of salivary flow in 123 subjects aged 12–49 and found that it was not related to height, age, pulse rate, or systolic or diastolic blood pressure. Since the same authors

found a direct relationship between rate of flow and pH, it might be inferred that these factors are not related to the pH of saliva. Deakins *et al.* (17) also found no correlation between rate of flow and the age of their subjects. Wenger (58) found that, for 488 aviation students, salivary pH had no correlation greater than .15 with any of 24 other physiological measures except salivary output; however, in a similar study of 201 *Operational Fatigue* patients, Wenger (58) found correlations of .28 with palmar conductance, −.20 with sublingual temperature, and −.20 with O_2 consumption. In a factor analysis of this latter study, salivary pH showed the highest correlation (.48) with a factor expressing the autonomic nervous system.

"Organic" Dysfunction

Stengel (53) determined colorimetrically the pH of saliva in patients with different gastric disorders and found that the pH was not significantly different in subjects with gastric hyper- or hypoacidity. Reindel (42) tested the pH of 50 subjects colorimetrically. In 88 per cent of the cases a "normal" gastric acidity was accompanied by a "normal" salivary pH. Of 22 cases of hypoacidity, 77 per cent exhibited a decreased salivary pH. In the group with hyperacidity, 73 per cent had an increased salivary pH. This would indicate an inverse relationship between salivary and gastric acidity in patients with gastric disorders; however, Wenger (58) found that the salivary pH of 21 peptic ulcer patients did not vary significantly from a group of 488 aviation cadets.

Mayr (37) found that the salivary pH of 25 psoriasis patients ranged between 6.9 and 8.4, with most of the determinations on the alkaline side, while the range for eczema patients was 6.3 to 7.4, with most of the subjects approaching normality. Wenger (58) found no significant deviations in salivary pH in 16 asthma patients.

Behavior

Among the earliest studies of the changes of salivary pH with behavior was Starr's (51) investigation of four subjects under the influence of emotional stimuli. Secreted saliva tended to become more alkaline during emotional behavior, with an average increase of about 0.4 pH. In the same report he notes that 228 normal subjects had a mean salivary pH of 6.8, while 10 "hyperexcitable psychopathic" stammerers had a mean salivary pH of 7.7. Starr relates this finding to the increased depth and rate of breathing, since 200 specimens from 58 "sub-breathing" stammerers showed a mean pH of 6.0. Hafford (26), who used an electric pH meter, failed to corroborate Starr's findings. She found no

significant difference between the acidity of 19 stutterers and normals. The discrepancy in her results and Starr's might have been due to her use of more rigidly controlled experimental conditions and a more precise measuring device.

Rich (44) correlated the salivary pH obtained electrometrically from 39 undergraduate students, with personality ratings made by five of their fraternity brothers, and the salivary pH of 18 graduate students, with personality ratings made by each other and by two of their instructors. The correlations were:

	Good Naturedness	*Perseverance*	*Leadership*	*Aggressiveness*	*Excitability*
39 Undergraduates	.03	.07	−.09	.09	.28
18 Graduates	.09	.24	.10	.51	.45

Rich concludes, with Starr, that the more alkaline is the saliva, the more excitable is the individual. Wenger (58), however, found no significant correlation between salivary pH and the 13 personality factors of the Guilford and the Guilford-Martin scales. His samples of 325 and 197 aviation cadets were rather highly selected and do not preclude that sigificant relationships may exist in an unselected sample.

Winsor and Korchin (60), in a carefully controlled study with an electric pH meter, found that anxiety, fear, and states of depression are accompanied by a decreased salivary pH, and this change is directly related to salivary flow. Since the rate of secretion is easier to determine than acidity, they suggest that the former is a more useful indicator of autonomic balance. Wenger (58), in a study of 225 *Operational Fatigue* patients, in which pH measurements were taken with nitrazene paper, obtained a difference (C.R., 2.50) between the salivary pH of *Operational Fatigue* patients and a group of 488 aviation cadets. The patients had the more acid saliva. Their mean value was 7.1, while that of the cadets was 7.2. Further, a group of 98 patients classified as *Mixed Psychoneurotics* and *Anxiety Neurosis* demonstrated a mean of 6.9 and also differed significantly from the cadet control group. In a later study, Wenger, Furchtgott, and Voas (59), using a glass electrode pH meter, found that a group of 43 ambulatory hospitalized patients from a psychoneurotic ward had a mean salivary pH of 6.61, while a control group of 69 college students had a mean salivary pH of 6.45. The C.R. of 1.9 approaches significance at the .05 level. The mean pH for nine patient with diagnosis of *Mixed* or *Severe Psychoneurosis* was 6.61, for seven patients with diagnosis of *Anxiety Neurosis* it also was 6.61, for seven patients with diagnosis of *Hysteria* it was 6.62, for nine patients diagnosed as *Psychopathic Personality* it was 6.86, and for a

group of 11 miscellaneous cases consisting of epileptics, schizophrenics, and other psychotics it was 6.40. It should also be pointed out that of 32 patients listed above in the first four categories, 23 had salivary pH's which were 1 S.D. above or below the mean of the control group.

Wittkower and Pilz (61) examined 27 persons under hypnosis and 16 in the waking state and found that the various affective states increased or decreased secretion, depending on the individual rather than on the conditions of the stimulating situation. The kind of emotional state (fear, joy, disgust, etc.) had no specific influence on the pH changes produced. The determinations were colorimetric.

From the studies of this section it may be seen that some experimenters report a positive correlation between pH and excitability or disturbed states, others report a negative correlation, and still others find no consistent relationship. No more definite conclusions may be drawn at this time.

Summary and Conclusions

The recent literature on salivary pH has been reviewed because of the value that the measure may have as an indicator of (*a*) autonomic functions, (*b*) acid-base balance of the fluid matrix, and (*c*) certain overt behavior patterns.

It was found that the pH of mixed, unstimulated human saliva is between 6.44 and 6.85, with the different glands making different contributions to those values.

The following conclusions concerning methodology were reached:

1. The most adequate measuring device is the glass electrode pH meter, although for some large groups of subjects, colorimetric techniques do not seem to displace the mean values significantly from those obtained electrometrically.
2. Prevention of CO_2 loss from collected saliva does not seem to be as important a factor as was previously believed, providing the salivary pH is measured within five minutes after collection. Toluene is relatively ineffective in preventing CO_2 loss.

The following conclusions were reached in regard to the effect of the experimental conditions:

1. It is necessary to control or correct for the effects of:
   a) seasonal variations and particularly air temperature at the relatively high or low ranges,
   b) diurnal activity cycles,
   c) the sporadic occurrence of exercise, and
   d) the time interval since last ingestion of food.
2. Dietary differences have no apparent effect upon salivary pH.
3. When chemical substances are placed in the oral cavity to stimulate

the flow of saliva, the resulting pH is apparently inversely related to the pH of the stimulating object.

4. Saliva collected under conditions requiring jaw movements results in an increased pH.

5. Stimulation of the salivary glands by intravenous injection of autonomic nervous system stimulants or electric current may produce differential pH effects, depending upon the stimulating agent used and the gland stimulated or from which the saliva is collected.

The following conclusions were reached regarding the relationships among salivary pH and other physiological variables:

1. In mixed and parotid human saliva, salivary pH is directly proportional to rate of flow. There is some evidence to the effect that there may be differences in this respect between glands and/or species.

2. Results from subhuman animal experiments indicate that salivary pH does not reflect the acid-base balance of the blood unless great and/or prolonged changes occur in the blood. Contraindicating this conclusion is the fact that in fatigue and during exercise salivary pH decreases as the total blood acid increases.

3. Salivary pH is not related to the age and height of the subject and with the exception of salivary output yields only small correlations with any one of some 24 other physiological measures. It does, however, correlate highly (+.48) with Wenger's factor representing the autonomic nervous system.

4. There are some indications of a relationship between salivary pH and gastric acidity as well as certain skin disorders, but the data are inconclusive.

The following conclusions were reached regarding the relationships between salivary pH and behavioral variables:

1. Salivary pH differs among normal control groups, *Operational Fatigue* patients, and several groups of psychoneurotic patients. This and other evidence points to an "alkalinity-excitability" hypothesis, but the data are not conclusive, and more precise delineations of the experimental conditions are necessary if salivary pH is to be a useful index of various behavioral states.

BIBLIOGRAPHY

1. Baker, K. H., & Eye, M. G. The reliability of the pH of human mixed saliva as an indicator of physiological changes accompanying behavior. *Amer. J. Psychol.*, 1935, **47**, 222–240.
2. Baxter, H. Variations in the inorganic constitutents of mixed and parotid gland saliva activated by reflex stimulation of the dog. *J. biol. Chem.*, 1933, **102**, 203–217.
3. Berg, M., Mayne, A., & Petersen, W. F. Variability of blood pH and its association with meteorological factors. *Amer. J. Physiol.*, 1940, **130**, 9–21.
4. Benetato, G., & Oprisiu, C. Sur les variations du pH et des bicarbonates salivaires dans les états d'acidose et d'alcalose provoqués. *C. R. Soc. Biol. Paris*, 1938, **128**, 113–116.
5. Bienka, E. J. Über Änderung in der Menge der Chloride und der festen Bestandteile sowie über die Speichelsecretion der Submaxillardrüse in

Abhängigkeit von der Art und Stärke des Reizes. *Acta Biol. exp. Varsovie*, 1938, **12**, 93–95.

6. BIENKA, E. J., & SCZEPANSKY, C. Influence de la nature et de la force de l'excitant sur la quantité et sur la teneur en chlorures et le pH de la sécrétion salivaire. *C. R. Soc. Biol. Paris*, 1936, **123**, 32–34.
7. BILLMANN, E. L'électrode à quinhydrone et ses applications. *Bull. Soc. chim. Fr.*, 1927, **41**, 213–280.
8. BRASSFIELD, C. R. Flow and pH changes of submaxillary saliva associated with variations in acid-base equilibrium. *Amer. J. Physiol.*, 1945, **144**, 43–49.
9. BRASSFIELD, C. R., & HANDS, A. P. Changes in pH and the rate of flow of saliva accompanying pH changes in arterial blood during acetylcholine stimulation of the submaxillary gland. *Amer. J. Physiol.*, 1941, **133**, 222.
10. BRASSFIELD, C. R., HANSEN, E. T., & GESELL, R. Acid effects of ammonium compounds. *Fed. Proc. Amer. Soc. exp. Biol.*, 1946, **5**, 10.
11. BRASSFIELD, C. R., & HONG, C. J. Changes in pH and rate of flow accompanying pH changes in arterial blood during chorda tympani stimulation and during pilocarpine stimulation. *Amer. J. Physiol.*, 1939, **126**, 442.
12. BRAWLEY, R. E. Studies of the pH of normal resting saliva. I. Variations with age and sex. *J. dent. Res.*, 1935, **15**, 55–77.
13. BRONK, D. W., & GESELL, R. Electrical conductivity, electrical potential and pH concentration measurements on the submaxillary gland of the dog recorded with continuous photographic methods. *Amer. J. Physiol.*, 1926, **77**, 570–589.
14. CARLSON, V. R., & MCKINSTRY, M. L. Some studies on saliva with particular reference to induced acidity and alkalinity. *Dent. Cosmos*, 1924, **66**, 840–849.
15. CLAFF, C. L., & SWENSON, O. Micro glass electrode technique for determination of hydrogen ion activity of blood and other biological fluids. *J. biol. Chem.*, 1944, **152**, 519–522.
16. CLARK, G. W., SHELL, J. S., JOSEPHSON, J. B., & STOCKLE, M. E. Influence of diet upon inorganic constituents of human saliva. *Dent. Cosmos*, 1927, **69**, 500–513.
17. DEAKINS, M., CHEYNE, V. D., BIBBY, B. G., & VAN KESTEREN, M. Significance of salivary analysis upon small groups of subjects. *J. dent. Res.*, 1940, **20**, 161–170.
18. DE BEER, E. J., & WILSON, D. W. The inorganic composition of the parotid saliva of the dog and its relationship to the composition of the serum. *J. biol. Chem.*, 1932, **95**, 671–685.
19. EISENBRANDT, L. L. Variations in the pH of saliva of five individuals. *J. dent. Res.*, 1943, **22**, 147–156.
20. EISENBRANDT, L. L. Studies on the pH of saliva. *J. dent. Res.*, 1944, **23**, 373–374.
21. GANS, L. R. Experimental study in salivary reactions. *J. Amer. dent. Ass.*, 1926, **13**, 222–225.
22. GASSKILL, H. V. A method for preserving small samples of saliva for subsequent pH determination. *J. gen. Psychol.*, 1933, **9**, 256–257.
23. GLAUBIGER, A., & GALK, K. G. The determination of pH values of biologic body fluids. *J. lab. clin. Med.*, 1940, **4**, 369–376.
24. GREGERSEN, M. I., & INGALLS, E. N. The influence of rate of secretion on the concentration of K and Na in dogs' submaxillary saliva. *Amer. J. Physiol.*, 1931, **98**, 441–446.
25. GROSSMAN, L. I., & BRICKMANN, B. M. Comparison of diurnal and noctural pH values of saliva. *J. dent. Res.*, 1937, **16**, 179–182.
26. HAFFORD, J. A comparative study

of the salivary pH of the normal speaker and stutterer. *J. Speech Disorders*, 1941, 6, 173–183.

27. HENDERSON, M., & MILLET, J. A. P. On the hydrogen ion determination of normal saliva. *J. biol. Chem.*, 1927, 75, 559–566.
28. HOLZLÖHNER, E., & AIRAPETIANZ, E. Die Drüsentätigkeit bei Nervenreizung. III. Die Wirkung der Sympathikusreizung auf die Glandula submaxillaris und ihr Einfluss auf den Chordaeffect. *Z. Biol.*, 1933, 93, 491–506.
29. HUG, E., & MARENZI, A. D. Réactions paradoxales de quelques sécrétions après une injection endoveineuse d'acide. *C. R. Soc. Biol. Paris*, 1928, 99, 240–241.
30. KESZTYÜS, L., & MARTIN, J. Über den Einfluss von Chorda- und Sympathikusreizung auf die Zusammensetzung des Submaxillarspeichels. *Pflüg. arch. ges. Physiol.*, 1937, 239, 408–418.
31. KESZTYÜS, L., & MARTIN, J. [Physiochemical constants of submaxillary saliva obtained by chorda and sympathetic stimulation.] *Magyar Orvosi Arch.*, 1939, 40, 346–353.
32. KORCHIN, B., & WINSOR, A. L. The relationship of certain organic factors to individual differences in human parotid secretory rate. *J. exp. Psychol.*, 1940, 27, 192–194.
33. KRASSUSKY, V. K., KRYMSKAYA, M. L., & KOTLYAREVSKAYA, E. J. [Several features of the parotid gland secretion in ruminants.] *Fiziol. Zh. S.S.S.R.*, 1940, 28, 372–383 (*Biol. Abstr.*, 1943, 17, 8380).
34. LASHLEY, K. S. Reflex secretion of the human parotid gland. *J. exp. Psychol.*, 1916, 1, 461–493.
35. LEUNG, S. W. Changes in the salivary volume and pH with changes in oxygen content of inspired air. *J. dent. Res.*, 1946, 25, 170.
36. MCCLELLAND, J. R. The influence of various stimuli upon human saliva. *Amer. J. Physiol.*, 1922, 63, 127–141.
37. MAYR, J. K. Zur Pathologie der Psoriasis vulgaris. *Arch. Dermatol. Syph.*, 1930, 159, 598–604.
38. MOORE, A. U., & MARCUSE, F. L. Salivary, cardiac and motor indices of conditioning in two sows. *J. comp. Psychol.*, 1945, 38, 1–16.
39. NEKRASOV, P. A., & KHRANILOVA, N. V. [Effect of physical work on the unconditioned salivary reflex in man.] *Arkh. biol. Nauk*, 1934, 34, 593–604 (*Biol. Abstr.*, 1937, 11, 11074).
40. NIMS, L. F. Glass electrodes and apparatus for direct recording of pH in vivo. *Yale J. Biol. Med.*, 1938, 10, 241–246.
41. POPOW, N. A., & KUDRJAWZEW, W. Über den Einfluss der Reaction der Erregerung der Speichelabsonderung auf die aktive Reaction des Speichels. *Pflüg. arch. ges. Physiol.*, 1930, 224, 69–74.
42. REINDEL, W. Über die Gesetzmässige Gegensätzlichkeit zwischen den Acidverhältnissen von Mundspeichel und Magensaft. *Klin. Wschr.*, 1940, 19, 390–392.
43. RICH, G. J. The reaction of human mixed saliva. *Quart. J. exp. Physiol.*, 1927, 17, 53–56.
44. RICH, G. J. A biochemical approach to the study of personality. *J. abnorm. soc. Psychol.*, 1928, 23, 158–175.
45. SCHMIDT-NIELSEN, B. Microdetermination of pH in saliva. *Acta Physiol., Scand.*, 1946, 11, 97–103.
46. SCHMIDT-NIELSEN, B. The pH in parotid and mandibular saliva. *Acta Physiol., Scand.*, 1946, 11, 104–110.
47. SMITH, G. H. Factors affecting the deposition of dental calculus. *Aust. J. exp. biol. med. Sci.*, 1930, 7, 45–77.
48. SOYENKOFF, B. C., & HINCH, C. F. The measurement of pH and acid neutralizing power of saliva. *J. biol. Chem.*, 1939, 109, 467–475.

49. STARR, H. E. Studies of human mixed saliva. I. The determination of the hydrogen ion concentration of human mixed saliva. *J. biol. Chem.*, 1922, **54**, 43–54.

50. STARR, H. E. Studies of human mixed saliva. II. Variations in the hydrogen ion concentration of human mixed saliva. *J. biol. Chem.*, 1922, **54**, 55–64.

51. STARR, H. E. The hydrogen ion concentration of the mixed saliva considered as an index of fatigue and emotional excitation and applied to a study of the metabolic etiology of stammering. *Amer. J. Psychol.*, 1922, **33**, 394–418.

52. STAVRAKY, G. W. Mechanism of synergistic action of pilocarpine and adrenaline on salivary secretion. *Rev. Canad. Biol.*, 1942, **1**, 64–71.

53. STENGEL, F. Vergleichende Untersuchung über die Wasserstoffienenkonzentration (pH-Wert) des Speichels und des Magensaftes bei Magenkranken. *Arch. Verdauungs Kr.*, 1938, **63**, 191–195.

54. STEPHAN, R. M. Changes in hydrogen ion concentration on tooth surface. *J. Amer. dent. Ass.*, 1940, **27**, 718–723.

55. WAINWRIGHT, W. W., & BECKS, H. Human saliva. XVII. Is there a seasonal influence on rate of flow and Ca and P content of resting saliva? *J. dent. Res.*, 1946, **25**, 285–291.

56. WESSINGER, G. D. Comparative measurements of salivary pH. *J. dent. Res.*, 1947, **26**, 123–127.

57. WENGER, M. A. Seasonal variation in some physiologic variables. *J. lab. clin. Med.*, 1943, **28**, 1101–1108.

58. WENGER, M. A. Studies of autonomic balance in Army Air Forces personnel. *Comp. Psychol. Monogr.*, **19**, No. 4.

59. WENGER, M. A., FURCHTGOTT, E., & VOAS, R. B. A study of salivary output and pH in several abnormal groups. To be published.

60. WINSOR, A. L., & KORCHIN, B. The effect of different types of stimulation upon the pH of human parotid secretion. *J. exp. Psychol.*, 1938, **23**, 62–79.

61. WITTKOWER, E., & PILZ, W. Über affektive-somatische Veränderungen; zur affektiven Beeinflussbarkeit der Speichelsekretion. *Klin. Wschr.*, 1932, **11**, 718–719.

62. ZOETHOUT, W. D., & TUTTLE, W. W. *Textbook of physiology.* St. Louis: C. V. Mosby Co., 1943.

Received July 1, 1950.

RESEARCH WITH THE WECHSLER-BELLEVUE TEST: 1945–1950[1]

ALBERT I. RABIN AND WILSON H. GUERTIN
Michigan State College

Since the 1945 review (111) of work done with the Wechsler-Bellevue test and since its supplementation (136), a vast number of studies have been reported in the literature. Some of them are of major significance, while others have "nibbled" at various problems using the test as one of their main instruments. These past six years have seen a veritable flood of published material with numerous, frequently contradictory conclusions regarding the test itself and regarding its use in health and disease. Partial reviews and summaries concerned with its application to specific problems such as the study of personality (1), old age (95), and scatter (100) have appeared. However, no over-all review and stock-taking account of the whole gamut of applications of the Wechsler-Bellevue test has been undertaken. It is the purpose of the present review to give a general summary of recent research with the W-B[2] test. An effort has been made to cover the literature and investigations in which the test figures prominently. Our concern is with the original scale and not with other forms that have been published subsequently.

Our attempt will, in many instances, question the fruitfulness of the research and the conclusiveness of the findings. We hope that this critical examination will help prevent duplication and repetition of experimentation in certain barren and overworked fields and that it will point out some worth-while areas of investigation.

The first part of the present review will deal with the W-B (Wechsler-Bellevue) as a test of general intelligence, for which purpose it was originally designed. Test modifications, validation, and shorter forms will be discussed under this general heading. The second part covers studies of scatter and subtest patterning in which the main concern is the diagnosis of various and sundry emotional, psychiatric, and neurological conditions.

[1] Through September, 1950.

[2] In addition to the Wechsler-Bellevue abbreviation, the names of the subtests will appear in abbreviated form throughout the paper. The single letters *I*, *C*, *A*, *D*, *S*, and *V* stand for the verbal subtests: Information, Comprehension, Arithmetic, Digits, Similarities, and Vocabulary, respectively. The two letter combinations *PA*, *PC*, *OA*, *BD*, and *DS* correspond to the following subtests: Picture Arrangement, Picture Completion, Object Assembly, Block Design, and Digit Symbol. *FS*, *VS*, and *PS* stand for Full Scale, Verbal Scale, and Performance Scale, respectively.

The Wechsler-Bellevue as a Test of General Intelligence

Studies of Reliability

To date, the number of studies dealing with the reliability of the W-B as a psychometric instrument is rather limited. Many investigations deal with retest findings after therapeutic intervention. This material logically belongs to the second part of the present review, dealing with diagnostic problems in psychiatric populations.

Normal subjects. The study by Derner, Aborn, and Canter (2) is the only one dealing with the reliability of the W-B in normal subjects. They administered the test to 158 normals distributed over a fairly wide intelligence range. Sixty subjects were retested one week after the original testing, 60, four weeks later, and the remaining 38, after a period of six months. The mean IQ's as well as mean subtest scores clearly showed practice effects upon retest. The differences were reported to be statistically significant. The practice effect is lessened as the period between tests is lengthened. Coefficients of correlation for the combined groups range from .62 to .88 for the subtests, and the highest coefficient, .90, was reported for the Full Scale. The timed subtests and those subtests involving simple learning showed the greatest practice effects. The authors conclude that, "particularly the Full Scale, can be considered diagnostically dependable with normal subjects" (2, p. 178).

Psychiatric subjects. A similar study reporting retest findings with 53 psychiatric patients was reported by Hamister (61). Thirty-three of the patients were re-examined after an interval of one week, while the remaining 20 were retested after one month. The mean IQ increase for the entire group was 10.25 (statistically significant). No statistically significant differences were obtained between the IQ's of the group retested at the end of one week as compared with the group re-examined at the end of one month. The test-retest correlations of the various subtests ranged from the .50's to the .90's. A different approach, i.e., testing the split-half reliabilities of W-B subtests, was reported by Gilhooly (52). Only four subtests (*I*, *C*, *S*, and *V*) were studied by means of results obtained from 122 neuropsychiatric patients. The reliability coefficients for the subtests were reported as follows: Information (*I*), .79; Comprehension (*C*), .56; Similarities (*S*), .77; and Vocabulary (*V*), .94. The values reported in this study are similar to those reported in the retest studies discussed above (2, 61).

Summary. Thus far, only one study concerning the reliability of the W-B with normals is available in the literature. Studies utilizing psychiatric subjects are too inconclusive because of the constantly changing condition of the patients. On the whole, the Full Scale W-B proved to be a reliable intelligence scale. The reliability of individual subtests varies considerably. Practice

effects upon retest have been definitely demonstrated. It has also been shown that the effects of practice depend upon recency of the re-examination.

Correlation with Other Tests

Wechsler-Bellevue II. Gibby (**49**) found good agreement between Forms I and II of the W-B with a group of 32 neurotics. The correlation between the *FS*'s was .87. The tests were not wholly comparable, however, since comparisons of subtests revealed differences. Comprehension measures something quite different on the two forms as shown by the low correlation of .20 between the subtests. Picture Arrangement was next lowest, with a correlation of .49, while high correlations of .87 were obtained for *PC* and *BD*. A recent study by Hays (**63**) is concerned with the agreement in patterning of the mental defective on the two forms of the W-B. He found that there are differences in four subtests which were significant at the 1 per cent level of confidence. He did not present correlations as Gibby did, but he found significantly different means for *A*, *V*, *PA*, and *OA* weighted scores on the two forms. Wechsler's profile for defectives was best supported by the W-B Form I.

These studies urge caution in interchanging tests in research studies where something more than an aggregate measure of intelligence is required. It would seem particularly important to perform further investigations of profile patterning on the two tests for the various psychiatric groups, as Hays has done for mental defectives.

Stanford-Binet. Gothberg (**56**) found the W-B to yield IQ's about 15 points higher than the 1917 revision of the Stanford-Binet (S-B). Correlations between the S-B and the three W-B scales ranged from .73 to .89 with a group of dull-normal subjects. Superior adolescents were studied by Goldfarb (**54**) with the 1937 S-B. He found Form L yielded IQ scores about five points higher than the W-B. Correlations were .80, .67, and .86 with the *VS*, *PS*, and *FS*, respectively. Sartain (**121**) also found the 1937 S-B to be significantly higher than the W-B. Correlations of the W-B *FS*, Otis S-A, S-B, and ACE with grade-point averages failed to reveal significant differences. Guertin (**57**) found the W-B *FS* to give scores approximately five IQ points higher than the 1937 S-B. Subjects were 15 to 22 years old and of high-grade defective to dull-normal intelligence. Kutash (**85**) used a similar population except that many subjects were older. He found the W-B *FS* was 11 points higher than the 1937 S-B, but he did not control age. He concluded that the magnitude of the differences between the tests was largely a function of this uncontrolled variable. This is hardly surprising, since Wechsler's standardization procedure makes this result inevitable. Reported correlations are attenuated by the uncontrolled age factor and, hence, are of little meaning. Stacey and Levin (**131**) made a comprehensive report of a comparison of S-B and W-B scores for several groups of defectives. Correlations are attenuated by a somewhat limited

range of talent. The brighter subjects had slightly lower W-B than S-B IQ's, while the duller ones had higher W-B IQ's. Tagiuri (**133**) made a comparison of S-B and W-B vocabularies, but the paper was not seen by the authors.

Integrating these studies, we see that the W-B gives somewhat higher IQ's than the 1937 S-B in the lower portion of the intelligence range, but lower IQ's in the upper range. The 1917 S-B is probably still lower than the 1937 revision throughout the whole intelligence range.

Other intelligence tests. Altus (**7**), in his comparison of the Army Wechsler Form B and the W-B, concludes that the latter is slightly more valid. This conclusion seems to be based upon the slight difference between the biserial correlations between acceptance-rejection and IQ's at an Army training center. These were .57 for the W-B and .55 for the Army form. Following Wechsler's point-scoring method, Lindner and Gurvitz (**94**) restandardized the Revised Beta Examination. The correlation between their revision of the Beta and the W-B was .92. In spite of the difference in test content it would appear that they measure essentially the same thing, since the correlation coefficient is of the order of magnitude of a reliability coefficient.

The California Test of Mental Maturity (CTMM) was compared with the W-B by Clark (**20**). The mean total IQ's were comparable for his 50 veterans. The correlation between IQ's was .81. While these tests appeared to be equivalent, a surprisingly large discrepancy was found between the two verbal and the two nonverbal IQ's. Hermann and Hackman (**64**) reported the results of testing two samples of veterans, one with the W-B and one with the CTMM. The IQ's were higher for the sample receiving the CTMM, but sampling biases, recognized by the authors, make conclusions impossible.

The Kent EGY, and W-B were compared by Wright, MacPhee, and Cummings (**143**). The obtained correlation of .51 is rather meaningless because of the extremely restricted range of intelligence. While the mean IQ's are comparable, the authors were astute enough to study the distributions of IQ's in more detail. They found that while dull normals were underestimated, the defectives were overestimated by the EGY. The authors concluded that the EGY is not a satisfactory measure of intelligence for their population. A short series of items was drawn from the EGY for study by Knott and Cannicott (**80**). The psychiatric population showed a correlation of .86 between the EGY and the W-B. Altus (**8**), also interested in a brief verbal measure of intelligence, selected 13 items of general information from a larger battery. The correlation of his battery of items with the W-B *FS* (.73) is rather low, but he seemed to feel that the general information items could be used as a quick verbal measure of general intelligence.

Wright (**144**) studied the Shipley-Hartford Retreat Scale over a wide

range of intelligence in a neuropsychiatric population. The correlation with the W-B was .77. Further study of this test by Garfield (41) utilized a narrow sample of intelligence, viz., borderline defectives. The subjects were not adequately discriminated by the Shipley-Hartford Scale and the author omits correlational data with the W-B from his report. Magaret and Simpson (99) mention incidentally that the correlation between the W-B and the Shipley-Hartford was .87 for 25 heterogeneous hospitalized cases. French and Hunt (40) compared the Navy-Northwestern Successive Matrices with the W-B *C-V-S* as the criterion of intelligence. The low correlations between these two measures suggest that the Matrices sample some ability or abilities not measured by the *C-V-S*.

Miscellaneous instruments. Diamond (31) presents a report of an interesting investigation of the vocational guidance value of the W-B test. The W-B subtests and various aptitude tests (O'Rourke Survey of Vocabulary, Minnesota Spatial Relations, Minnesota Clerical) were factor analyzed. The subtests most heavily weighted with the three factors are Linguistic: *I*, *C*, *S*; Clerical: *DS*, *D*, *A*; Spatial: *PC*, *OA*, *BD*. He concludes that the W-B is a less adequate measure of vocational aptitudes than other more conventional tests, but that the information from the W-B, when treated in this three-factor frame of reference, is of value to the guidance counselor.

Buck (17), who devised the ingenious H-T-P projective drawing test, compared his quantitative scoring with the W-B as a measure of intelligence. He expresses the belief that it may tap aspects of intelligence not usually sampled by conventional intelligence tests. He obtained fair correlations with the W-B, but his sample had an excessive range of talent. His failure to allow for intellectual deterioration with increasing age complicates the comparison. This would tend to attenuate his correlation coefficients and to make the W-B mean IQ too high. Sloan and Guertin (127) used the W-B as the criterion of intelligence in their study of the psychometric value of the H-T-P. The limited range of talent employed—borderlines and high-grade defectives—makes the reported correlations difficult to interpret. The authors found the H-T-P IQ to be 15 points higher than the W-B *FS* in their adolescent sample, and they question the adequacy of the H-T-P standardization.

Summary. It is rather obvious from the literature that the W-B is a well-accepted measure of intelligence. Furthermore, the standardization is equally well accepted, as demonstrated by the tendency of investigators to use W-B scores as criteria of intelligence in validating newer instruments. This is somewhat in contrast to the tendency seen in the early review (111) of checking W-B validity against other instruments such as the S-B. Only two studies (7, 121) reviewed here utilize criteria of intelligence other than psychometric tests. Neither study demonstrated the superiority of the W-B, yet both indicated a fairly high degree of validity for it. Only when the discrepancy between IQ's

is between the W-B and the S-B, which is based upon an even larger standardization group, would investigators consider the fallibility of the W-B standardization. The psychometrist's faith seems to be well placed in Wechsler's standardization and the test in general. Some question might be raised of the adequacy and meaning of standardization results at the low end of the intelligence distribution. There is evidence (29) that length of institutionalization influences performance on this test. Wechsler seems to have used some new admissions and some already committed to state institutions, but he does not specify the proportion of each. The test should be used for examination of cases similar to those used in the standardization sample to assure meaningful classifications and profile analyses.

Research-oriented clinicians should see a lesson to be learned from this review, namely, the use of a proper range of talent in correlation studies. Many studies (40, 41, 121, 127, 143) have shown rather inconclusive results because the samples were selected without regard for the distribution of intelligence in the *whole* population. Gothberg (56) is to be particularly commended for her recognition of this requirement and her description and comparison of her sample with Wechsler's standardization population. One comparison of test results (143) disclosed important information, concealed by an over-all comparison of means, through an analysis of the forms of the two distributions of IQ's. This methodological approach indicates the need for more thorough analyses of results than is obtained through the use of standard gross statistics.

Short Forms

Since the first short form (*C-A-S*) was introduced by Rabin, many other forms have appeared. The results of the earlier investigations of short forms were so encouraging that many investigators have published since the last review (111). All the modifications using only W-B subtests are summarized in Table I. McNemar's report (102) is too extensive to include in the table and is reviewed briefly in the body of this section.

Combinations of W-B subtests. Gurvitz (58) suggests using only two subtests, *D-PA*, which is a rather radical step from point of view of brevity. Despite the length of this form and the general feeling about the unreliability of *D*, he found a correlation of .90 with the *FS*, with a large group of subjects. Since this was higher than anticipated, he commendably made further investigations with different populations and found confirmation of this high correlation. Geil (17) proposes the use of a longer form, the *C-S-D-BD*, on the basis of a *FS* correlation of .97 with a sizable group of unselected prisoners. He felt that this form was especially useful in screening out organics and psychotics. In their study of various short forms Cohen and Steisel (24) mentioned the diagnostic value of Geil's form but reported that the highest validity was obtained for *C-A-S*. They do not confirm Gurvitz's findings (50) of higher validity for *D-PA* than for *C-A-S*, but their *N* was only 32. However, Springer's report (130) showed a high correlation of .96 for

TABLE I

VALIDITY COEFFICIENTS FOR VARIOUS SHORT FORMS

Form	*Investigator*	*N*	*Population*	*Criterion*	*r*
D-PA*	Gurvitz (58)	523	Heterogeneous	WBFS†	.90
D-PA	Patterson (107)	50	Hetero. psychiat.	WBFS	.81
D-PA	Cohen & Steisel (24)	32	Superiors	WBFS	.52
D-PA	Cotzin & Gallagher (26)	154	HGMD & bdrlne	WBFS	.75
C-A	Cummings *et al.* (27)	420	Bdrlne	WBFS	.93
C-A	Patterson (107)	50	Hetero. psychiat.	WBFS	.85
C-A	Hunt *et al.* (69)	528	Normals	NGCT‡	.58
C-A	Cotzin & Gallagher (26)	154	HGMD & bdrlne	WBFS	.53
S-PA	Hunt *et al.* (69)	528	Normals	NGCT	.72
C-A-S	Springer (130)	100	Bdrlne	WBFS	.96
C-A-S	Patterson (107)	50	Hetero. psychiat.	WBFS	.89
C-A-S	Cohen & Steisel (24)	32	Superiors	WBFS	.78
C-A-S	Hunt *et al.* (69)	528	Normals	NGCT	.78
C-A-S	Cotzin & Gallagher (26)	154	HGMD & bdrlne	WBFS	.59
C-V-DS	Patterson (107)	50	Hetero. psychiat.	WBFS	.93
C-V-DS	Patterson (109)	100	Normals	WBFS	.90
C-A-PA	Hunt *et al.* (69)	528	Normals	NGCT	.72
S-D-PA	Hunt *et al.* (69)	528	Normals	NGCT	.75
C-S-DS	Bay & Berks (11)	200	Court cases	WBFS	.93
C-S-D-BD	Geil (47)	250	Random prisoners	WBFS	.97
C-S-D-BD	Cohen & Steisel (24)	32	Superiors	WBFS	.87
C-S-D-BD	Patterson (107)	50	Hetero. psychiat.	WBFS	.93
C-S-D-BD	Cotzin & Gallagher (26)	154	HGMD & bdrlne	WBFS	.84
V-C-BD-PC	Patterson (107)	50	Hetero. psychiat.	WBFS	.96
V-C-BD-PC	Patterson (109)	100	Normals	WBFS	.96
I-PC-PA-DS	Geil (47)	250	Random prisoners	WBFS	.95
I-PC-PA-DS	Patterson (107)	50	Hetero. psychiat.	WBFS	.95
I-PC-PA-DS	Cotzin & Gallagher (26)	154	HGMD & bdrlne	WBFS	.74
V-I-BD-S	Kriegman & Hansen (83)	207	Bdrlne	WBFS	.95

* The letter abbreviations for the Wechsler-Bellevue subtests are as follows, in order of appearance in the table: *D* = Digits; *PA* = Picture Arrangement; *C* = Comprehension; *A* = Arithmetic; *S* = Similarities; *V* = Vocabulary; *DS* = Digit Symbol; *BD* = Block Design; *PC* = Picture Completion and *I* = Information.

† Wechsler-Bellevue Full Scale.

‡ Navy General Classification Test.

the *C-A-S* with 100 cases of suspected low mentality. This coefficient entails the fallacy of correlating the part with the whole, so Springer reports the correlation of the *C-A-S* with the W-B *FS* minus these three subtests. The obtained coefficient of .89 was still surprisingly high. Hunt, Klebanoff, and Mensh (69) utilized a criterion other than the W-B and found for a large group of normals a correlation of .78 between the *C-A-S* and the Navy General Classification Test (NGCT). This is a fairly respectable degree of agreement between the two tests.

Cummings, MacPhee, and Wright (27) found a correlation of .93 between their proposed *C-A* form and the W-B *VS* for a large number of Naval recruits of questionable mental ability. Still another modification was proposed by Bay and Berks (11), the *C-S-DS*. They found a correlation of .93 with the *FS* for a mixed sample of court cases and college students. They felt that this form is valuable not only because of its high degree of validity, but also because it retains much of the diagnostic value of the *FS*. Patterson (107) was interested in several combinations of subtests proposed by previous investigators (27, 47, 58, 130). On a group of 50 variously diagnosed subjects he confirmed the high correlations of the short forms with the *FS* found by the original investigators. He concluded that vocabulary should be included and proposed the *V-C-BD-PC* and *V-C-DS*, which have correlations with the *FS* of .96 and .93, respectively. He feels that many of the diagnostic features of the W-B are retained in these batteries. In a later study (109) he found very similar degrees of relationship between his two suggested short forms and the *FS*. Patterson concludes that the selection of a short form for use with a given subject should depend upon the level of intelligence being studied. His *V-C-DS* proved best for superior subjects (109). Cotzin and Gallagher (26) obtained good results with the *C-S-D-BD*, first proposed by Geil (47), when applied to mental defectives. This form was found to correlate .84 with the *FS*, while others (27, 58, 107, 130) proved less valuable with this population. Kriegman and Hansen (83) studied still another combination of subtests, *V-I-S-BD*, and found a correlation of .95 with the *FS* for a large group of borderlines.

McNemar (102) has evolved a statistical technique for arriving at correlations between various combinations of subtests and the *FS* for Wechsler's standardization data. This economical process made it possible to study many combinations. He reports the ten forms with highest correlations with *FS* for pairs, triads, quartets, and quintets of subtests. His correlations are high, but lower than those generally reported in the literature. He attributes this to the use of too broad a range of talent in many studies. However, it is the impression of the present reviewers that there was more frequently a restricted range of talent operating in the studies surveyed here. His enthusiasm for the use of Wechsler's standardization sample is understandable, since it guarantees a proper range of talent, but the use of a nonstandardization type of population is more justifiable if the investigator is concerned with developing a short form for use with that special type of case, as is so often the need. Cohen (23) reported the application of a short form to four different psychiatric groups, but the paper was not available for review.

Brief forms using some W-B subtests. Hunt, Klebanoff, and Williams (69) incorporated the Thurstone modification of the Stanford-Binet

Form L vocabulary in an attempt to retain this subtest in a more practical form. This study is of special interest since it is the only recent one that has employed a criterion other than the W-B *FS*. Their *C-V-S* showed the highest correlation with their criterion, the Navy General Classification Test (NGCT), among all the forms which they have studied. Their large sample of normals gave a correlation of .86 between the tests. This tendency to find substitutes for W-B subtests in the short forms also appears in a recent study by Hunt, French, and Mensh (**70**), in which they propose the possibility of substituting the McGill Verbal Situation Test in place of *C* or *S* in *C-V-S*. Williams, Hunt, and French (**141**) present normative data for their *C-V-S* based upon 1,000 subjects.

Summary. While most of the studies have been chiefly interested in validity correlations, the agreements in IQ's with the *FS* seem to be at a high level. The attention to correlation is pragmatically justifiable since a restandardization or correction for the short form can be made if the relationship to the *FS* warrants it. Necessary corrections or methods of deriving IQ's from the short forms other than simple prorating have been proposed by some of the investigators (**27**, **83**, **141**).

The correlations cited here are surprisingly high in view of the attenuating conditions that seem to exist. The frequently encountered restricted range of talent (small IQ range) and the diminished reliability of psychiatric populations are the cases in point. On the other hand, conditions are often present which tend to raise the coefficients spuriously. Correlating the part with the whole (short form vs. *FS*) tends to raise the coefficient artificially, although only .07 in Springer's study (**130**). Occasionally, combinations of subtests have been selected to take advantage of sampling irregularities. This has not been a serious error, since large *N*'s have usually been employed and the rechecking of results by later investigators with independent samples has confirmed the original reports. This need to avoid capitalizing on sample irregularities through the use of more than one sample was recognized by several of the investigators (**27**, **58**, **69**, **109**).

All studies of short forms agree in several of their conclusions. It is clear that most short forms are sufficiently valid in their measurement of the same thing that the W-B measures, to be used with borderlines with considerable time economy. The application of short forms to superiors is not so well established, but there is some evidence (**24**, **109**) that the proper combination of subtests will give meaningful results. It would seem desirable to select the proper short form for the intelligence level one is concerned with. Franklin (**38**) has shown that caution must be used in applying short forms to populations other than those on which they were standardized. With delinquent Negro boys the short form investigated is more than four IQ points lower than the W-B *FS* IQ. Unfortunately no correlations were reported.

The studies reviewed here often make the point that no short form retains the full diagnostic value of the *FS*. Some (**11**, **47**, **107**) have tried to select the subtests that would provide the most valuable diagnostic clues, but all agree that there is no substitute for the *FS* in a complete case study.

The intensive research in this area has been facilitated by the availability of routine W-B data. Often one gets the idea that the aim of the study reported was to "get another publication" rather than to make a systematic contribution. Since so many short forms have been proposed and validity demonstrated, it would appear that further research in this area should be confined to systematic attacks such as McNemar's recent contribution (**102**). However, there still seems to be a need for a combination of subtests which will yield valid results on a variety of different populations.

Applications with Special Populations

Superiors. The application of the W-B to populations other than the normal standardization population is a very practical problem. Estes (**33**) points out the difficulties encountered in a scatter analysis of the test results of superior adults. The excessively high verbal and vocabulary weighted scores obtained by these subjects introduces an artifact if scatter is measured from vocabulary. In his study, negative vocabulary scatter on *PC* and *DS* are regarded as a consequence both of using this excessively high vocabulary as a reference point, and the ceiling limitations of the subtests.

Mental defectives. The validity of Wechsler's mental defective subtest pattern for institutionalized mental defectives was investigated and confirmed by Cutts and Sloan (**28**). Cotzin and Gallagher (**26**) presented mean weighted scores in their article which confirm the above investigation. With institutionalized defectives, *A* is extremely low, while *OA* is highest, and, in line with this, the *P* IQ is higher than the *V* IQ. Cutts and O'Kelley (**29**) worked on the problem of the effect of institutionalization on the subtest distribution of defectives. They found that the *VS* was depressed; lowest scores were on *I*, *D*, and *A*. They conclude that their patients with long hospitalization fail to develop verbal abilities to the same extent as other defectives.

Aged. Madonick and Solomon (**98**) were interested in W-B findings with the aged. They gave the test to 50 residents of a denominational home and a home caring for formerly successful executives and professionals. In spite of the fact that their sample appears to be biased in the direction of high IQ's, the authors felt the need to explain the high mean IQ's obtained (*V* IQ, 117.1, and *P* IQ, 115.6). They suggest that the nature of the test construction or the environment, which was stimulating because of a communality of interests, may account for the obtained results. A similar study by Chesrow, Wasika, and Reinitz (**19**) utilized only 20 white males. The order of weighted subtest scores was very similar to that found by Madonick and Solomon. They also found that the vocabulary IQ was usually higher than the *FS* IQ. The *V* IQ's and *P* IQ's were comparable.

Minority groups. A comprehensive analysis of the discrimination of the subtests at various levels among mentally-subnormal Negroes was

made by MacPhee, Wright, and Cummings (**104**). They found different patterns at different intellectual levels, i.e., the subtests contributed differently to the *FS* score at the different levels. At the low levels, *S* contributed most to the obtained *FS* score, while as the intelligence increased, all tests came to contribute rather equally. On the basis of subtest correlations with the W-B *FS* they were able to conclude that the W-B is a reliable measure for the southern Negro. Franklin (**38**) also found that the subtests, with few exceptions, discriminated effectively between classifications of intellectual level of delinquent Negro boys.

Hodgson (**67**) studied the test performance of delinquent Mexican boys and found that *DS* failed to discriminate between borderlines and defectives, while *S* failed to discriminate between normals and dull-normals. The order of magnitudes of subtest correlations with the whole test was similar to that obtained with native white Americans. Glaser (**53**) investigated the intelligence of Jewish immigrants. Through an ingenious control and alternate directions in Yiddish and German, he demonstrated the cultural difficulties encountered with *I* and *PA*. Digits was harder for the immigrants, probably because of language handicap and anxieties. Surprisingly, the *V* IQ was somewhat higher than the *P* IQ. The thesis is that performance items (*PA* and *PC*) can prove to be greater obstacles for the immigrant than verbal items because of their cultural loading.

Miscellaneous populations. A unique approach to evaluating intelligence in the blind with the verbal scale of the W-B is seen in a report by Levine and Blackburn (**87**). They studied a large number of W-B verbal scales of newly blinded soldiers. Army General Classification Test scores on these men before their blindness occurred show them to be a representative sample of soldiers with respect to intelligence. The correlation between the prior test and the postblindness W-B verbal scale was .83. The authors found that, unlike individuals who had been blinded for a long time, these cases were significantly lower in *D*. This is attributed to inattention and anxiety. They conclude that as far as their sample is concerned, no impairment of intellectual function can be observed as a consequence of blindness.

Naval recruits, under observation for intellectual deficit, were discriminated well by intelligence levels, according to a report by Lewinski (**89**). Each of the verbal subtests "discriminated effectively," according to the author. Schmidt (**123**) examined the intelligence levels of prisoners with the W-B *FS* and group tests. While the group tests indicated subnormal intelligence, the W-B contradicted this. Lewinski (**93**) also reported an analysis of W-B results on 1,000 soldiers. While the report is primarily a consideration of vocabulary, Wechsler's standardization receives support since all three scales have similar means (near 100) and there is only a very slight correlation between CA and IQ.

Summary. The application of the W-B has not been restricted merely to the standardization population. Application to special groups has been of value when the investigators have taken the differences in subjects into account. The internal validation criterion (subtest vs. whole test) has proved a ready means of attack on these problems. Investigations utilizing this method (**38, 67, 89, 104**) might serve as the models for validity studies of the W-B applied to other special samples.

Refinements and Critiques

Order of subtests. According to Wechsler (**138**) the subtests have their items placed in approximate order of difficulty. Practicing clinicians have noticed that some items seem to be misplaced. A paper by Rabin, Davis, and Sanderson (**112**) reports the order of difficulty for the items in six subtests, based upon the performance of 300 normals and confirmed upon 1,000 psychiatric cases. Some differences from Wechsler's orders of difficulty were found for these populations. Jastak (**72**) has recently reported a similar study with a large group of psychiatric cases. He demonstrated that differences in the kinds of subjects in the sample influence the order of difficulty of the subtests. Some of the sex differences in order of difficulty are especially interesting. In general, his results are quite at variance with Wechsler's data, and he finds that the orders of difficulty are more similar among his various subgroups than between Wechsler's orders and any of those for his subgroups or combined subjects. Certain advantages to an even progression are mentioned in the paper: (1) economy and ease of administration, (2) setting reasonable limits of individual testing, and (3) deriving clinically meaningful measures of intratest scatter.

Sorsby (**128**) presented a revised order of difficulty for vocabulary in connection with a simplified administration procedure. The basal and ceiling levels are established on the basis of only three consecutive successes and failures. He found that these middle 60 per cent of the words were as discriminating as the full range and reduced administration and scoring time by about 50 per cent.

Klugman (**79**) was interested in the effect of order of subtests on digits score. He found that the mean digit score was highest when the subtest came in the middle of the battery. Since the score was higher at the beginning than at the end of the battery, it would seem that fatigue is a more important factor in reducing efficiency on digit recall than is initial anxiety associated with adjustment to the test situation. No psychometric advantage to placement in the middle is demonstrated by this study, and it is of theoretical interest only. No matter where a subtest is placed, it is standardized in that position, so that any constant decrement related to subtest order becomes unimportant. Had the author studied the variances under the three conditions of placement within the battery, he might have been able to attribute

greater variance to increased unreliability, which would have practical importance.

Administration procedure. Sappenfield (120) recommends the use of cardboard cutouts or templates, which lift up and leave the *OA* pieces in their proper positions for administration. Derner and Aborn (30) make a similar suggestion, but use a board with inserts for the *OA* pieces. They are inserted upside down, and a sliding cover prevents them from falling out. The board is placed upside down before the subject, and the pieces drop on the table in their proper places as the cover is slid back. This procedure prevents the subject from seeing the pieces before he is ready to begin the assembly task. These procedures would have their greatest usefulness with neophytes.

Scoring. A practical suggestion by Welsh (139) is the combination of time and accuracy scores in one table. This would reduce computational errors in combining the two credits and avoid the possible omission of time credit. Rubinstein (119) proposed a new method of accurately recording *BD* performance. The procedure is too detailed to be adopted by the average clinician but could be useful in complete case studies and in research.

Analysis. Foster (36) pointed out an error inherent in most clinical evaluations based upon W-B weighted subtest scores. Since Wechsler's weighted scores are based exclusively upon the standardization results of the 20–24 year olds, they can not be interpreted directly when applied to older age groups. Foster suggests that tables for weighted scores for corresponding age groups be included in the manual in order to make meaningful comparisons of subtest weighted scores possible. Barnett (10) elaborates on this idea and suggests that the tables give some indication of the importance of the deviation of the obtained score from the mean score for that CA group. This has been done by this author in his tables of deviation-scores in sigma units. Such tables should be of considerable value to research workers who are studying scatter and profile analysis.

Gilhooly (50, 51) tested the hypothesis that mean intratest scatter vs. mean score is a negatively accelerated function in the average to superior intelligence range. Only a slight curvilinear relationship was found with his group of 122 neuropsychiatric cases. The *eta* throughout the whole intelligence range was only .25 and was not statistically significant.

Miscellaneous studies. Gerstein (48) presented an interesting procedure for analyzing vocabulary responses. She utilizes Yacorzynski's principle that the "right answers" on *V* are not prima facie evidence of adequate knowledge of the word meaning. Her discussion suggests avenues for further research that might disclose more information relating to the nature of verbal performance on this test.

Spache (129) found positive correlations between the quality of the

UNIVERSITY OF MICHIGAN LIBRARIES

responses to *S* and *C* and the weighted scores for the subtests. Some correlation would, of course, be expected since the quality of the response contributes to the weighted score.

A total of 2,000 reports by 50 psychologists were examined by Luchins (97). The practices which evoke criticism are many but special emphasis is placed upon the incorrect application of Terman's norms to W-B results. A critique of W-B weighted scores is offered by Brown (16). He discusses the "ceiling-floor" effect operating in the testing of subjects with very high or very low ability. He points out that pattern analysis, scatter, and deterioration measures are inappropriately applied when a subject's ability is beyond the ceiling or floor of the test. This is really a discussion of a generally recognized psychometric principle which is not always observed. It is unfair to imply that the ceiling-floor limitations are the fault of a test. The only legitimate criticism of a test is its narrow range of applicability. In this respect, it would appear that the W-B provides as wide a range of measurement as is practical. It can be objected that Wechsler's norm tables include IQ values beyond the desirable limits, which might prove misleading to naive testers who believe that the listing of IQ values is sufficient sanction for using them.

Summary. Some of the suggested modifications in W-B procedure encountered in this review appear to make contributions meriting immediate adoption. Others appear to be of minor importance and are justifiably omitted from standard procedures. It would seem particularly important to publish tables of deviation scores for each CA level for scatter evaluation (10), omit the low tabular IQ values in the conversion tables in Wechsler's tables, and to adopt a new vocabulary procedure (128).

The Wechsler-Bellevue as an Aid in Psychiatric Diagnosis

Retest Findings after Therapeutic Intervention

In an earlier section of the present review, some retest studies with the W-B have been considered. These were retest studies per se without the introduction of known modifying factors in the general condition of the subjects. The following paragraphs will deal with retest studies following fairly radical therapeutic intervention.

Shock treatments. Kessler's (77) 20 schizophrenic patients were retested with the W-B following 9–25 electric shock treatments. The interval between tests ranged from one to five months. With the exception of three cases, all patients showed IQ increases, the mean increase for the group being 10.6. Though increases were noted in all subtests, statistically significant changes were noted only in *PC* and *C* and the three IQ's obtainable with the W-B. High correspondence (95 per cent) between increases in IQ and psychiatrists' judgment of improvement

was also noted. A less detailed and less well-controlled study by Dubin and Thaler (32) reports retest findings on 11 schizophrenics following an unspecified number of insulin and electric shock treatments. The interval between tests ranged between three and 20 weeks. The mean increase in IQ's is reported to be 17 points. All but two subtests (*A* and *V*) showed a statistically significant increase upon retest. However, no relationship between IQ increase and psychiatric condition was obtained. A similar conclusion regarding the lack of relationship between IQ changes and changes in psychiatric conditions was reached by Rabin (113) in ten electrically treated patients. He also found that a rise in IQ is positively correlated with the length of time that elapsed between the last treatment and the retest. Patients retested within a week following the last treatment showed decrements or very small increments in IQ. Those retested from 28–60 days following the last treatment showed a uniformly high increase in IQ. The results are inconclusive due to a possible high sampling error. Further experimental investigation in this area is needed. Fisher (34) employed W-B Forms I and II, alternately, with a group of electric-shock-treated "ambulatory depressed patients." Little change occurred in the unimproved group, but the improved group showed significant changes in *C*, *S*, *DS*, and the three IQ's. It may be noted that this study has two shortcomings: (1) The equivalence of the two forms of the W-B may be doubted (49), and (2) the possible differential loading of diagnostic categories in the improved and unimproved groups is a factor unaccounted for. Carp (18) reported on retest findings in 42 insulin-treated schizophrenics. The mean interval between tests was 73 days. The retest mean IQ was significantly higher than the one obtained during the first testing situation; 76 per cent of the patients showing a gain in IQ, 19 per cent showed a loss, and the remaining five per cent showed no change. The IQ changes were attributed by the author to retesting with the same instrument, rather than to the effects of the therapy. The loss in 19 per cent of the cases, however, remains unexplained. Like Dubin and Thaler (32) and Rabin (113), Carp was unable to show any relationship between improvement or lack of it, and changes in IQ. Five W-B subtests were part of a battery used by Luborsky (96) in examining 12 psychiatric patients before, during, and after a course of electric shock treatments. The five schizophrenic patients included in the group tended to show a drop in subtest weighted scores during the course of treatment. Following the final treatment a rise in scores was general. No measures of significance were reported.

Prefrontal lobotomy. Retest findings on ten patients who underwent prefrontal lobotomy operations were reported by McCullough (101). The patients were examined one week before the operation, and again eight weeks following it. The only statistically significant gains were noted on *D* and *PA*. Most resistant to change were the *V*, *I*, and *S*

UNIVERSITY OF MICHIGAN LIBRARIES

subtests. A single case study (145) reported a loss of 17 IQ points following the operation.

Summary. Retest studies following therapeutic intervention raise considerable doubt whether the demonstrated gains in W-B scores and IQ's are due to the therapy rather than to practice effects. The study with normals (2) strengthens this doubt. Results regarding correspondence between rise in IQ level and improved psychiatric conditions are equivocal. Effects of recency of treatment upon the retest findings bear further investigation. Differential effects of therapeutic intervention upon the several subtest scores have not been clearly established.

The Problem of "Scatter" and Its Ramifications

As was indicated in the earlier review (111), the W-B test gave the problem of a scatter a new lease on life and controversy. The results obtained with the Binet were negative or inconclusive at best. The earlier review discussed a number of attempts to use test patterns in diagnosis. At best, group differentiations were attained, but individual diagnosis via subtest configuration still remained only a desideratum.

The Menninger study. At about the time of publication of the last review, there appeared a very extensive and intensive study of the use of scatter in differential diagnosis by Rapaport and his collaborators at the Menninger Clinic (114). In more than three hundred pages the W-B was described, and a detailed presentation of several measures of scatter in 271 subjects (19 diagnostic subgroups and one "control" group consisting of 54 patrolmen) was given. The three main measures of scatter used in this project were:

1. *Vocabulary scatter*, consisting of differences between the vocabulary score and scores on the other subtests.
2. *Mean scatter*, consisting of "the difference between any subtest score and the average of all the subtest scores" (excepting *D* and *A*).
3. *Modified mean scatter*, consisting of the differences between a verbal or performance subtest score and the mean of all the other scores of their respective scales.

The task performed by Rapaport and his collaborators is a gigantic one. However, as far as the adequacy of experimental design and conclusiveness of findings are concerned, considerable doubt is cast upon this opus by reviewers and other investigators trying to check their results. A summary of the more pertinent criticisms of the Rapaport study, without reference to extraneous stylistic aspects (15), is as follows:

1. The "control" group is really not a control group (46). Its inadequacy for comparison with pathological groups stems from failure to control sex, age, education, cultural and socioeconomic status, etc. Age is probably most important in this criticism since scores on the W-B, like those on any other intelligence scale, are subject to variation with the age of the subjects.

2. The 19 clinical groups resulting from a breakdown of 217 cases is quite inadequate, since the number in each group becomes much too small for proper statistical comparison and manipulation (**103**).

3. Another criticism, which is also valid in connection with previous scatter studies, is that the scatter measures may be rough indicators of differences *between groups*, but are not valid or useful in individual diagnosis (**46**).

4. For adequate differential dignosis the whole gamut of psychopathological conditions should have been included (**84**). It would have been especially important to include several categories such as organics, psychopaths, and feebleminded, and to differentiate them, respectively, from schizophrenics, neurotics, and a variety of deteriorated conditions.

5. It also seems doubtful whether the use of such statistical techniques as the *t*-test and chi-square on this type of data is justifiable (**46**).

6. The intermingling of clinical observations with the data and conclusions that are primarily to be drawn from quantification findings is very confusing and emphasizes the lack of validity of the results (**3, 46**).

It appears that the most valuable contribution made in this study by the Menninger group lies in the realm of qualitative description of the W-B subtests and in the hypotheses made regarding the possible psychological processes involved in these performances. Krugman (**84**) states it well in his review: "Using the data as a springboard, Rapaport has employed a combination of unusual clinical insight, acute analytical treatment and profound theoretical reasoning, to synthesize a diagnostic approach . . . " (p. 285). This statement, probably minus the glowing adjectives, summarizes well the nature of the contribution.

From the viewpoint of quantitative research and experimentation this study has failed to produce any clear-cut or conclusive findings concerning scatter. On the contrary, it has added to the already existing confusion of contradictory findings on various groups which differ in age and in control of the several pertinent factors involved. Some of the studies quoted in the latter part of the paper will have a further, direct bearing upon the methods and conclusions of the Menninger group.

Other measures of scatter. Pascal and Zeaman (**106**) have empirically derived an index of W-B scatter which differs somewhat from previous attempts. It is based on differences between three pairs of subtests (*I-C*; *BD-DS*; *V-PC*) and the absolute difference between *S* and *BD*. The added resulting figures comprise an "index," the application of which shows significant differences between normals and neurotics, and between neurotics and psychotics. Moreover, they conclude that their index reflects "the progressive seriousness of psychological disturbances" and that it gives "an estimate of level of adjustment." Absence of psychological rationale and group, rather than individual, differentiation are the major criticisms of this and many similar approaches. The index relationship to diagnosis may suffer from the same shortcomings of another empirically derived "ratio" criticized by Webb (**137**).

UNIVERSITY OF MICHIGAN LIBRARIES

Holzberg and Deane (**68**) have suggested a measure of "intratest scatter" instead of the conventional intertest scatter within the W-B scales. The scatter of the items within some subtests was reported by the authors to distinguish to a significant degree between organics, neurotics, and schizophrenics. The "altitude" approach to the measurement of scatter was suggested recently by Whiteman (**140**). This measure utilizes the top scores of subtests, rather than their mean, as a reference point for deviation and scatter analysis. Statistically significant differences between normals and schizophrenics were reported as a result of the application of this type of scatter.

Scatter in differential diagnosis. Using Rapaport's data, Wittenborn (**142**) tested several neatly arranged hypotheses which are important in profile or scatter diagnosis. The hypotheses may be roughly stated as follows:

1. All patients (Rapaport's) having the same diagnosis have a similar W-B test score pattern.
2. Subtest correlations between pairs of patients of the same diagnostic group are higher than those between pairs of different diagnosis.
3. Correlation coefficients between pairs of patients of the same diagnosis are higher than coefficients obtained "between pairs of hypothetical patients for whom two or three of the subtests have the same rank" (p. 437).

The method of rank order correlation was used in the investigation of each of these hypotheses. The coefficients of correlation (in terms of *rho*) were so small that they cannot be considered as support for the hypotheses. The author concludes that for Rapaport's groups a scatter pattern "is of no greater value as a psychometric supplement to diagnosis than an inspection of the subtests for a few conspicuous instances of failure and relative success" (p. 439).

In his review of scatter analysis, Mayman (**100**) was rather optimistic about its value in differential diagnosis. However, subsequent work did not seem to justify entirely this optimism. Patterson (**108**) reported 78 per cent agreement between psychologists' diagnoses based on scatter analysis and that of psychiatric diagnoses is open to several major criticisms. The "blind" diagnoses were not actually blind since the majority were seen by the psychologist who diagnosed them; the test reports were a factor in establishing the psychiatric diagnoses. Moreover, test patterns (scatter) on which the diagnoses were based were not reported. A close examination of the actual comparison of "agreements" by Garfield (**42**) shows that the level of time agreement is really much lower than the 78 per cent reported. A study of similar design by Teicher and Singer (**134**) reported 52 per cent agreement between scatter and psychiatric diagnosis. Another 8 per cent showed agreement to some extent. Test patterns of 50 neurotics and 50 schizophrenics, rigidly controlled with respect to a number of significant fac-

tors (age, sex, education, etc.), were reported by Kogan (**82**). In this study pairs of patients having similar W-B patterning were compared with each other, regardless of diagnosis. The author's conclusion is that his "investigation does not support the commonly held conception that the WB patterns are a useful tool in the differential diagnosis of functional disorders . . . " (p. 44). Similar results, prompting "a note of caution," were also reported by Peixotto (**110**). Inadequacy of finer differentiation within the normal range has also been reported (**35**).

Summary. It is quite possible that some of the high correlations between psychiatric and W-B diagnoses obtained by different investigators are due to the inspectional techniques of the clinician rather than the numerical scatter values. One may also agree with the reviewers (**1**) who state that "Score-patterns successfully differentiate groups, but are of comparatively uncertain value in the diagnosis of individuals" (p. 108). The efficacy of W-B scatter in *individual* or differential diagnoses has not been demonstrated. Intratest scatter (**68**) and "altitude" (**140**) as a reference point in scatter may merit further investigation.

Indices of Deterioration

A more particular aspect of scatter that has been reported extensively in the literature is that of deterioration. In the third edition of *Measurement of Adult Intelligence*, Wechsler (**138**) proposes an index of deterioration, which he calls the "Mental Deterioration Index" (MDI).

Diagnosis of organicity. Levi and his collaborators (**86**) found this method useful in differential diagnosis between psychogenic and organic aspects of disorders in the 15 cases described in their clinical study. However, these investigators seemed fully aware of the inconclusiveness of their findings. They left the following points to be clarified by future investigations:

1. Determination of a critical percentage of loss indicative of definite deterioration.
2. The extent to which psychogenic disorders may simulate organic conditions.
3. The extent to which test deterioration is unaccompanied by physical symptomatology.

The test pattern and the application of the MDI to 50 brain-injured patients were studied by Allen (**3, 4**). The results shows that the *I* subtest is most resistant to the effects of brain injury and that the *OA* is one of those most affected. Moreover, the application of the MDI disclosed that only 54 per cent of the brain-injured were identified by it (using 20 per cent as the "definite" indication). Allen also suggested a modified index consisting of the difference between the sums of the weighted scores of two pairs of subtests (*I* and *C* minus *D* and *DS*). His suggestion was that a difference of 5 points or more between these pairs of subtests would be a more definite index of deterioration.

UNIVERSITY OF MICHIGAN LIBRARIES

Blake and McCarty (12) compared Allen's data on the brain-injured with results from 50 college students. Using the 10 per cent level as critical, they discovered that 14 patients and 12 normals were misdiagnosed. Thus, according to Wechsler's method, 26 per cent of the total group of 100 subjects were misdiagnosed. When they applied Allen's modification to the normals, 17 cases, or 34 per cent, were misdiagnosed. The authors admit that all they have shown is some statistical significance for the MDI; however, they caution that " . . . this report is not concerned with the efficiency of this score in differential diagnosis" (p. 418).

Recent work by Kass (76) essentially shows agreement with Allen's (4, 5, 6) findings. He concludes that in two-thirds of his cases the MDI failed to detect or confirm cases of brain damage.

Deterioration versus mental deficiency. Boehm and Sarason (13) chose to investigate whether the formula offered by Wechsler is successful in distinguishing intellectual deterioration from mental deficiency. Twenty-two "familial" mental defectives ranging in age from 15–63 were the subjects of this study. In all instances the neurological examination was negative, and feeblemindedness was evident from early childhood. The mean deterioration loss was 15 per cent for the entire group. Fourteen of the cases showed losses from 22 to 48 per cent. The authors, therefore, quite correctly conclude: "The deterioration index . . . cannot be considered as a valid tool for distinguishing between mental deficiency and deterioration" (p. 358).

A study by Sloan (125) lends added support to the conclusions of Boehm and Sarason (13). This investigation applied Wechsler's MDI to 80 male mental defectives, predominantly of the familial variety. Using the 20 per cent criterion, nearly 70 per cent of this population would be considered as deteriorated. However, neither clinically nor psychometrically could such a contention be supported. Fortunately, intelligence test findings were available on 56 cases that had been examined during previous years. A re-examination did not produce any significant differences. The coefficient resulting from a correlation between the magnitude of the difference and the deterioration index was positive, but too small to be considered statistically significant. Thus, Wechsler's MDI was shown again to be inadequate in differentiating "deterioration" from mental deficiency.

Age and deterioration. Hunt's analysis (71) of Wechsler's original standardization data, in which the subtest score means of the young, nondeteriorated group (ages 20–24) are compared with those of an old group (55–59), showed that only two of the subtests seem to "hold up" with age (*I* and *C*). Moreover, the author also shows that *BD* and *DS* are the most vulnerable subtests. He consequently suggests that a ratio based on these pairs of subtests might be most sensitive and distinguishing.

Copple's analysis (25) of 1,403 records of a wide age range (20–69) awards "some validity" to Wechsler's MDI. The author feels, however, that the use of his conversion table for obtaining a senescent decline score, based on a multiple regression equation, is preferable. A study referred to earlier (19) supports the use of the index in the aged and shows a positive correlation with CA.

Johnson (73) correlated the MDI with CA for sizable groups of epileptics and feebleminded. The coefficients of correlation obtained were quite high for two epileptic groups (.60 and .46) and quite low (.22) for the feebleminded. Yet, the correlation between the length of time the patients had had seizures and the MDI is reported to be low by the same investigator. Thus, the conclusion one is tempted to draw from the first part of the study—that there is a relationship between age in epileptics and the MDI, which can probably be attributed to the history of seizures—is not supported in the second part. Since there is considerable difference in the reported mean ages of the groups, the results are difficult to interpret.

Differentiation between various groups. A comparison between psychoneurosis and schizophrenia was made by Schlosser and Kantor (122). Their population of 163 subjects consisted of 56 psychoneurotics, 65 paranoid schizophrenics, and 42 mixed schizophrenics. They did not use all of the subtests in the scale. The "hold" subtests used in the ratio were: *I*, *V*, *PC*, and *OA*. The "don't hold" subtests were: *D*, *A*, *BD*, and *DS*. A priori, the expectancy would be that the schizophrenic group, particularly the nonparanoid, would show a higher loss than the psychoneurotics. However, no statistically significant differences were obtained between the three groups. It appears that when the deterioration index is used, psychoneurotics, paranoid, and mixed schizophrenics show similar levels of loss of efficiency. The inadequacy of *I*, *C*, and *V* as indicators of the level of intelligence in the premorbid state of patients was pointed out by Rappaport and Webb (115).

The MDI was investigated by Rogers (117) with a total of 349 subjects, subdivided into seven groups, some described in the literature and others obtained from counseling and mental hygiene services. The findings indicate that normals differ significantly from the clinical groups. However, the maladjusted groups do not differ significantly from each other when the MDI is used. The index does not seem to differentiate adequately between organics and one group of patients with functional disorder. Results similar to those reported by Blake and McCarty (12) were obtained by Rogers (118) when Allen's index (4) was used.

Comparison with other measures of deterioration. Garfield and Fey (45) studied the relationship between Wechsler's measure of deterioration and that obtained from the Shipley-Hartford. The 100 hospitalized males revealed an agreement of .37 between these two measures.

UNIVERSITY OF MICHIGAN LIBRARIES

Magaret and Simpson (99) report a correlation of only .10 between these two indices of deterioration. Furthermore, they failed to find evidence for validity of either method of estimating deterioration. The clever procedure of correlating psychometric decline over a period of illness with measures of deterioration, and a comparison of the indices with psychiatrists' ratings of deterioration gave negative results.

Fox and Birren (37) compared Wechsler's MDI with findings in the Babcock-Levy test. The correlation obtained between the two indices in 50 senescent normals was not significantly different from zero. The need for independent validating criteria was emphasized by them.

Summary. Wechsler's original remarks (138), based on empirical results and not entirely borne out by his own statistics, set off a wave of studies, the results of which are predominantly negative. The original discussion of the deterioration index was chiefly in connection with the "normal process of deterioration" due to age. However, many investigators and clinicians did not lose time in jumping the gap from old age to organicity. It may well be that the reduction of mental functioning due to old age is qualitatively different from brain trauma or brain disease. Moreover, does all brain trauma or disease of necessity produce marked deterioration? Are not the foci of pathology and the pretraumatic and predisease personalities important factors to be considered before all of these cases are tossed into the "organic hopper?"

A more serious indictment of the MDI comes from the studies on familial mental deficiency. They raise the question whether the "index" is only characteristic of old age, or whether it may not be the result of a pattern to be found in a variety of conditions not related to old age. Another difficulty is the inadequacy of the MDI in differentiating "organics" from some other functional groups. Further statistical refinement of the index followed up by its application to a variety of comparable (age, IQ, etc.) clinical populations may produce more definitive results concerning its value as a clinical tool in differential diagnosis.

Effect of Organic Involvement

An attempt to measure changes in the intellectual functioning due to physiological or somatic variations and due to structural brain insult has been made by several investigators who utilized the W-B scales. In these instances, the investigations deal with effects noted in the entire scale or the subtests in it, but not with any specific "deterioration index."

Differential diagnosis. One of the studies which falls in the category described above was made by a group of investigators at the Boston Psychopathic Hospital (55). Over a 13-month period, 428 patients were examined by means of the W-B scales. Of these, 92 were considered as showing "patterns" of brain damage. These judgments were based on eight subtests of the W-B scale (exclusive of *OA*, *DS*, and *V*). When checked by clinical examination (history and a variety of neurological

tests) 72 per cent of the cases were verified by clinical diagnosis; in 10 per cent the diagnosis was inconclusive, and in the remaining 18 per cent no clinical evidence of brain damage was obtained. Seventeen per cent of 40 hospitalized patients "with no psychological abnormalities to suggest brain damage" had some abnormal clinical findings. The study suggests numerous valuable *qualitative* observations based on the work of Goldstein and his collaborators. However, most of these observations are not, strictly speaking, the W-B test. The authors suggest such qualitative signs as irritability, mood changes, irregularity in speed of performance, stickiness in shifting, circumlocution, slurring, and many others as indications of organicity. It may be questioned to what extent these signs overlap with the "mental examination" on which the clinical diagnosis is based at least in part. The inspection of the figures of subtest distribution yields a "semiquantitative" conclusion that the patients with brain damage "tended to show relatively great preservation of ability on comprehension and information items. . . . Considerable loss . . . on *D*, *A*, and all the nonverbal items" (55, p. 172). This conclusion is based on a general impression rather than on detailed statistical analysis and tests of significance. One cannot agree that these findings "establish" the W-B as a reliable instrument in the diagnosis of brain damage. All that can be said is that some psychologists acquainted with clinical signs of organicity may be able to make a diagnosis of brain damage while using the W-B test.

An interesting approach to the diagnosis of cerebral pathology with the W-B test was presented by Hewson (65). This worker used as many as 13 ratios between various combinations of subtests in a comparison of patients with cerebral pathology (concussion cases, cerebral arteriosclerosis, CNS syphilis, etc.) with a variety of psychoneurotics and a group of normals. These ratios apparently have been empirically derived. The author states that the obtained quotients of the normal group were distributed between those of the neurotics and the cerebral pathology group. However, an examination of whatever meager quantitative data are presented does not justify the statement. It is also interesting to note that this worker advocates combinations of ratios for adequate diagnosis. A great deal of raw data on the 59 neurotics and 98 cerebral pathology cases are given. In addition, some results on normal subjects and on patients having a variety of additional diagnoses were presented. The relationship of these data to the efficacy of the ratios in diagnosis is not clear. For some of the numerical data presented, the indications are that the overlapping between organics, normals, and neurotics is considerable when the combination of the several ratios is used. The connection between some of the speculations and alleged conclusions and the data presented is very nebulous and hardly comprehensible to the present reviewers.

This procedure was also applied by Guttman (59) in comparing 30

UNIVERSITY OF MICHIGAN LIBRARIES

organics with 30 normals. Sixty per cent of the organics and 17 per cent of the normals were diagnosed as organic. The author considers Hewson's method superior to other types of W-B indices of organicity.

Epilepsy and migraine. Lewinski's studies on epilepsy (92) and migraine (91) may also be included in this section, though the organic basis for both of these disorders is far from having been established. These two disorders show similar rank ordering of subtests. The *rho* coefficient of correlation between the rank order of subtests in epilepsy and migraine is .936—almost unity. The *rho* when epilepsy is compared with neurosis is lower (.718+.11) but statistically significant. However, no significant correlation between normal and organic subtest ranks was obtained. The group findings show that performance was particularly good on *OA*, *BD*, *PC*, and *C*.

Single case reports. An interesting speculation regarding organicity in test performance was brought forth in Cofer's study (22) of a case of postinfectious encephalopathy in whom definite neurological involvement was found. The W-B (exclusive of *A*, *DS*, and *PA*) was administered on three occasions with a seven- and a two-week interval. The second administration was under hyoscine. Under the effects of the drug, the total IQ rose 14 points. The rise in the performance IQ was much more marked than was the verbal IQ. A lapse to the previous level of performance took place two weeks after the treatment was terminated. Though this is a single case, it is reasonable to question whether the deficit which may be observed in such cases via test performance is stable and unidirectional. Possibly other factors than the brain damage were operating.

Kogan's case report (81) of a meningo-encephalitic patient indicates highest subtest scores on *I* and *A* and particular difficulty with *C* and *BC*. Thinking, as a result of qualitative findings from the W-B and other tests, was considered to be extremely "concrete." These results are fairly consistent with Cofer's findings (22) regarding the rise in the performance tests under hyoscine.

Summary. It may be stated that the studies reviewed in this section add little to our systematic understanding of intellectual changes in organic disorder. The variety of patterns in "organics" raises the question whether the structural insults of different localizations may produce differential effects upon intellectual functioning. The possibility exists that some common elements observed in the patterns may be due to a similar *reaction of the personality* to the damage. It is unfortunate that many of the studies do not report all of the scores for the W-B subtests, thus making precise comparison of findings regarding intra-individual patterning very difficult, if not impossible.

Schizophrenic Patterning

Considerable impetus to subsequent research in schizophrenia was given by the studies reviewed in 1945 (111) and by the publication of

Wechsler's revised edition of *Measurement of Adult Intelligence* (**138**). Wechsler expanded his discussion of diagnostic indications in the test.

Studies utilizing Wechsler's diagnostic indicators. Brecher (**12**) obtained from Wechsler's discussion as many as 14 diagnostic "signs" for schizophrenia, 12 for organics, 13 for neurosis, 13 for psychopathic personality, and 12 for mental deficiency. Her study of a group of 40 schizophrenics indicates the futility of these signs. If the patients who were diagnosed clinically as schizophrenic were to be diagnosed on the basis of Wechsler's signs of schizophrenia as opposed to those characteristic of other disorders, only 32.5 per cent of the cases would be correctly diagnosed.

Olch (**105**) compared test results on 32 schizophrenics with those of two age groups (17–29 and 40–49) of Wechsler's standardization population. A number of characteristics of the schizophrenic pattern suggested by Wechsler and other investigators mentioned in our earlier review (**111**) were critically examined. Olch found no significant differences between the verbal and performance IQ's of schizophrenics; neither did she find that the age factor has any effect on these conclusions. However, she did find significant differences between younger schizophrenics and normals in individual subtest scores. These subtests are: *DS*, *C*, *S*, and *PC*. A finding substantiating earlier studies is the greater variability in test performance of schizophrenics as compared with Wechsler's normals.

In Levine's study (**88**), five judges were asked to diagnose W-B records on the basis of Wechsler's discussion in the recent revision of his book (**138**). The records were those of 110 white males distributed among a large number of diagnostic groupings. Among these, 46 cases were diagnosed as schizophrenic. The findings indicate that: "the agreement between WB diagnoses is of borderline statistical significance. The percentage of agreement ranges from 59.7 to 67.2 for the cases classified" (**88**, p. 29). Despite the fair statistical reliability for the ratings, the disagreement in diagnosing schizophrenics vs. nonschizophrenics ranged from 17–32 per cent, and there was overlapping of the profiles and unreliable discrimination in some cases. Thus, the separation of schizophrenics from other patients on the basis of W-B patterns is only somewhat better than chance, but considerably better than the results obtained by Brecher (**14**) who employed the "sign counting" method in examining their efficacy for ***individual*** diagnosis.

Johnson (**74**) obtained his score patterns by averaging the individual's subtest deviations from the mean of his subtest scores. Agreement of observed subtest deviation with that proposed by Wechsler ranges from 8 per cent on *OA* to 87 per cent on *S*. Variability was greater than reported by Wechsler, sex differences were not found, and a positive relationship between variability and intellectual level was obtained.

UNIVERSITY OF MICHIGAN LIBRARIES

Other schizophrenic patterns. Garfield (43), in his study of 67 schizophrenics and 46 "controls" (nonpsychotic psychiatric patients), utilized the deviation of each subject's subtest scores from his own mean of the subtests, excluding vocabulary. The data were reported in five degrees of deviation, two positive and two negative, clustering around an insignificant middle category. Table I shows the largest differences in the *BD* test in which the schizophrenics are favored, while the control may be, on the other hand, compared favorably with the schizophrenics on the *C* test. Though other trends were noticeable, there was "considerable overlapping between the schizophrenics and control groups." It also appears to this investigator that "no clear-cut patterns are discernible." The alleged higher verbal score than performance score in schizophrenics is criticized and receives no statistical support. Rabin's index, Rapaport's idea (96) that *BD* is less affected by schizophrenia than *PC*, and Wechsler's "tentative criteria which could be checked objectively" (*OA*, *BD*, *PA*+*PC*, *I*+*BD*) are also criticized and given no support by Garfield's results. The general conclusion of the author is that "there are no reliable scatter patterns on the Wechsler-Bellevue scale for schizophrenic subjects" (43, p. 35). In view of the inadequacy of the "control" sample (neurotics, psychopaths, alcoholics, etc.) it may be wondered whether such a categorical statement is warranted.

In a later study (44) the same investigator reported on a larger sample of 109 schizophrenic patients. In a comparison between the ranks of the means of subtests of his own group with those of Olch (105), Magaret, and Gilliland *et al.*, mentioned in the earlier review (111), he points out the inconsistency between group findings. The *DS* is about the only subtest that shows consistently low scores in the samples compared; other subtests, however, show variations of considerable magnitude. He feels, quite justifiably, that the assumption that different schizophrenic groups are comparable is not warranted, due to the differences in diagnostic criteria used. It also appears that the discrepancy between the verbal and performance scales (in favor of the former) is a function of educational and intellectual level. The difference is considerable in the educationally and intellectually superior group, while it is nonexistent in the average and subaverage groups. Variations in the rank order of the several subtests in the different groups were also noted. The author feels that in view of his findings, Rapaport's (114) relative neglect of such factors as age, IQ, and educational level in diagnostic studies of scatter is not justified. Finally, a tabulation of the percentages of subtest deviation shows wide variability for the entire schizophrenic sample and leads to the varification of the previous conclusion that there is no clear-cut schizophrenic test pattern.

Klein's (78) "application of the multiple regression principle" is illustrated in the study of Rapaport's (114) 72 schizophrenics and 54 normal controls. A comparison of these two groups on vocabulary

scatter, employed by Rapaport, yields some interesting results. "Schizophrenic impairment" is due least to *D* and *OA* and is most evident in *C*, *PC*, *A*, and *DS*. In attempting to differentiate *individuals*, the regression equation misidentified 26 per cent. The "linear combination of subtest scores yielded a multiple *R* of .45 with the criterion of group identity" and is significantly different from zero. Subtests *C*, *PC*, *DS*, and *A*, in order of decreasing effectiveness, were almost entirely responsible for "the discrimination capacity of the battery." The critique of the sampling method that may be applied to Rapaport's work is also applicable to this study.

Subtypes of schizophrenia. Freeman (39) compared 28 hebephrenic schizophrenics and 28 paranoids with the same number of normals matched for age, sex, race, educational background, and even religious affiliation. The data reveal that the discrepancy between the verbal and performance scale scores markedly favors the verbal score for the paranoid group, but is in the opposite direction in the hebephrenics. The total scores, verbal scores, and performance scores are highest for the normals and lowest for the hebephrenics. The hebephrenics function at a significantly lower level than do the controls on both scales, while the difference between them and the paranoids is only significant with respect to the *VS*. A significant difference between the paranoids and controls was found on the *PS*. The *FS* comparisons between the three groups were all statistically significant. The *FS* scatter and *PS* scatter did not show discrimination between the groups. The only significant difference on verbal scatter was found in the comparison of the hebephrenics and controls. The use of covariance did not alter the findings to any appreciable extent. Freeman concludes that the "scatter approach has not yet proved a valid one in the diagnostic testing of hebephrenic and paranoid schizophrenics."

An attempt to measure patterns in 245 schizophrenics was made by Harper (62) by means of Fisher's discriminant function. Reliable group differentiations between some major types of schizophrenia were obtained. However, the final conclusion is that the W-B ". . . is not a highly adequate tool for psychiatric diagnosis" (p. 295).

Summary. The findings concerning the use of the W-B in the diagnosis of schizophrenia are, in the main, negative. The results concerning the discrepancy between the *VS* and *PS* in schizophrenics remain inconclusive, since they differ from study to study. Individual differentiating subtests are not discriminating between schizophrenics and other groups. Several scatter approaches (mean scatter, vocabulary scatter, etc.) do not demonstrate the reliability of a specific schizophrenic pattern. Despite repeated warnings in the literature, many significant factors remain uncontrolled in many of the studies. Wechsler's "sign" approach did not prove fruitful in differential diagnosis. Judges utilizing Wechsler's diagnostic hints were somewhat more successful. Apparently, a combined approach in which some quantitative data are utilized along with "clinical insight" is most practical in the diagnostic process.

UNIVERSITY OF MICHIGAN LIBRARIES

Patterns in Neurosis

Anxiety. A scatter analysis (deviation of each subtest from the mean of all subtests) of the W-B records of 25 anxiety neurotics was reported by Lewinski (**90**). No characteristic pattern was obtained beyond a few qualitative observations. Rashkis and Welsh (**116**) devised 12 signs of anxiety which were based on W-B records of 30 cases. Seven of the "signs" are due to "temporary inefficiency" on *D*, *I*, *BD*, *A*, *OA*, *PC*, and *PA*, while the remaining are extraneous and qualitative in nature (apprehension, distractibility, somatic complaints, etc.). A comparison between the anxiety patients and a group of nonanxiety patients did not yield statistically significant differences in the incidence of the signs. The qualitative signs appeared among the ones that tended to differentiate between the groups most effectively. Shoben's comparison (**124**) of 18 anxious with 17 nonanxious college students yielded no significant differences in the number of signs listed by Rashkis and Welsh (**116**).

Warner's (**135**) study reported a comparison of W-B findings in 60 anxiety neurotics with those of 65 normal controls. Excepting the higher relative position of *PA*, *OA*, and *PC* in the anxiety neurotics, the investigator found that intertest variability, differences between verbal and performance IQ's, and relative *D* and *A* scores did not differentiate between the two groups.

A study in which a number of important factors were controlled in a comparison of 50 anxiety reaction cases with 25 orthopedic hospital cases, was reported by Heyer (**66**). A comparison of the groups with respect to mean weighted subtest scores, intertest variability, and subtest deviation from vocabulary yielded no positive results. *Group* results supported Wechsler's contention of *V* IQ higher than *P* IQ. However, this sign proved useless in *individual* diagnosis. No characteristic "pattern" of anxiety was produced on the W-B scales. A question regarding the adequacy of the control group may be raised. Orthopedic surgical patients frequently may have severe enough anxiety reactions to be comparable to those of anxiety neurotics.

Compulsive alcoholism. Similar negative conclusions have been reached by Halpern (**60**), who studied a group of 47 compulsive drinkers. There was no "characteristic organization of mental abilities which would serve to distinguish them either from normal subjects or from other clinical groups" (p. 472).

Summary. The results of the several studies do not support the contention of quantitative W-B findings characteristic of anxiety or compulsive trends. Some helpful clinical and qualitative observations may be made during the administration of the test, but they are not an integral part of it or uniquely evoked by it.

Psychopaths and Delinquents

Strother (132) studied the W-B's of 14 psychiatrically diagnosed psychopaths as a check on Wechsler's proposed pattern. Only one sign, low *A*, was in the expected direction, so the author concludes that these signs do not appear with sufficient consistency to be useful. Institutionalized behavior problem groups were examined by Clark (21). His Mexican and American soldier and civilian problem groups all showed depressed W-B *VS*, low *A*, high *OA*, and *PA*+*OA* greater than *PC*+*BD*. He concludes that Wechsler's psychopath pattern (138) is substantiated. Franklin (38) applied the W-B to delinquent Negro boys. Although aware of the fallacy of comparing their pattern with Wechsler's, which was obtained from a different type of sample, he found considerable agreement. Altus and Clark (9) studied a group of white and a group of Mexican institutionalized problem boys. Since four of Wechsler's adolescent psychopath signs were found, the authors concluded that his pattern was corroborated. Sloan and Cutts (126) studied institutionalized defective delinquents. They concluded that there is some agreement with both the Wechsler pattern for mental defectives and that for psychopaths. However, a control group of nondelinquents (28) showed just as many of the psychopath "signs."

Summary. The notorious lack of consistency in the use of the label means that different investigators study different subjects. The test results then are treated as group data, with means, which will obscure important individual differences. There is also the questionable application of the *adolescent* psychopath pattern of Wechsler to mature subjects. Lastly, it should be mentioned that in the original data (138, p. 214, No. 15) from which Wechsler derived his pattern, the "nonpsychopathic" delinquent controls showed a test pattern that was as similar to that of the "psychopaths" as it was to the standardization population. Considering all these difficulties, the reader will probably be surprised to see *any* agreement between investigators.

General Summary

The eleven years since the original publication of Wechsler's scale of "global intelligence" in 1939 have shown that his instrument is an easily adaptable and pliable tool for research purposes. Recent studies no longer question the validity of this test as a measure of intelligence (111). On the contrary, the vast popularity and wide usage of the test have made it a commonly used measuring rod for comparison and validation, if not actual calibration of newer and more recent techniques. Moreover, even parts of the scale or certain combinations of its subtests have been variously reported and advocated as rapid and time-

UNIVERSITY OF MICHIGAN LIBRARIES

saving devices in the assessment of intelligence. Many of the combinations show correlations of high order with the entire test, though their use must be limited to special situations. For the most part, findings are not readily comparable for they were obtained on samples which differ considerably from each other with respect to IQ range, education, age, sex, and other factors of varying significance.

The psychologist's rejection, in recent years, of his role as a "mere psychometrist" and his aspiration to diagnose the whole personality, rather than only its intellectual component, gave rise to a strong interest in the use of the various measures as means of psychiatric diagnosis. Underlying the flood of research that was stimulated by this interest is the hypothesis that the serveral intellectual factors measured by the test are differentially affected by a variety of emotional, psychiatric, and neurological conditions. Most intratest variability seemed to have been attributed to pathological conditions without sufficient regard for such variabilities in the test profile of the everyday, nonhospitalized person. The possibility that individual differences exist in the premorbid state to such a degree that they are mistakenly thought to be amplifications due to the morbid state has not been adequately considered. Thus, group tendencies of various nosological classifications have been frequently hailed as aids in psychiatric diagnosis. In fact, the dozens of studies published have not changed much the state of affairs in respect to "scatter," described six years ago (111). To be sure, many more groups have been tested and compared, and some additional measures of scatter contributed. Yet the scatter mountain gave birth to a mouse. It still is true, however, that " . . . the various measures of scatter and variability—the different patterns, have succeeded in differentiating *groups*, but not *individuals*" (111, p. 420). The empirically devised measures still point out "trends" and "tendencies" but cannot replace clinical observations. On the clinical and "qualitative" level some of the studies have brought forth certain observations and insights that may be of practical value. There is even some evidence of greater success in diagnosis of W-B records if one is not held strictly to some particular pattern, but pools the quantitative patterns with qualitative observations. Apart from the criticism that can be leveled against the multitude of investigations for failure to study large enough groups and to control numerous important factors, the difficulty is with the underlying hypothesis itself. Implied in the hypothesis is the notion that the criterion is clear-cut and of a homogeneous nature. But, alas, not all schizophrenics are alike or obtain the same diagnosis in different

hospitals; not all "organics" have the same amount and locus of brain destruction, etc., etc. The error is in assuming stability, uniformity, and homogeneity of the criterion. To do that is to build on sand.

Possibly some attention should be concentrated upon the measuring instrument itself (46, 75, 78). There are deficiencies in the structure of the test if it is to be used for the purposes just discussed. Inclusion of more significant functions and the increase in number of items within each measured function may lead to a closer approximation of the desired results. Also, experimentation with an improved scale, with subjects (preferably "normal") under varying conditions (stress, emotion, etc.) may assist in the refinement of intellectual measures which are particularly vulnerable to the more exaggerated and chronic conditions in psychopathology.

Almost any test employed by a skillful clinician and insightful worker may prove an "aid in differential diagnosis." However, so far, we have failed to demonstrate the effectiveness of certain quantitative formulas and test score variabilities as an aid in individual diagnosis. The theoretical contribution of such studies to the general field of psychology has been meager. It is doubtful that the continued clogging of the periodical literature with testimonials pro and con, based on scanty and uncontrolled evidence, will be of further benefit.

BIBLIOGRAPHY

1. ABELSON, H. H., & ELLIS, A. Other devices for investigating personality. *Rev. educ. Res.*, 1947, **17**, 101–109.
2. DERNER, G. F., ABORN, M., & CANTER, A. H. The reliability of the Wechsler-Bellevue subtests and scales. *J. consult. Psychol.*, 1950, **14**, 172–179.
3. ALEXANDER, FRANCES, CRUTCHLOW, EILEEN, & HOFFMAN, MARY. A selective survey of the Wechsler-Bellevue section of Rapaport's *Diagnostic psychological testing*. *Canad. J. Psychol.*, 1947, **1**, 111–115.
4. ALLEN, R. M. The test performance of the brain injured. *J. clin. Psychol.*, 1947, **3**, 225–230.
5. ALLEN, R. M. A note on the use of the Wechsler-Bellevue Scale Mental Deterioration Index with brain injured patients. *J. clin. Psychol.*, 1948, **4**, 88–89.
6. ALLEN, R. M. II. The test performance of the brain diseased. *J. clin. Psychol.*, 1948, **4**, 281–284.
7. ALTUS, W. D. The comparative validities of two tests of general aptitude in an Army Special Training Center. *J. appl. Psychol.*, 1946, **30**, 42–44.
8. ALTUS, W. D. The validity of an abbreviated information test used in the Army. *J. consult. Psychol.*, 1948, **12**, 270–275.
9. ALTUS, W. D., & CLARK, J. H. Subtest variation on the Wechsler-Bellevue for two institutionalized behavior problem groups. *J.*

UNIVERSITY OF MICHIGAN LIBRARIES

consult. Psychol., 1949, **13**, 444–447.

10. BARNETT, I. The use of *Z* scores in equating the Wechsler-Bellevue subtests. *J. clin. Psychol.*, 1950, **6**, 184–188.
11. BAY, MARGARET, & BERKS, MARY. Comprehension, similarities, and digit symbols of the Wechsler-Bellevue Scale used in a court clinic. *Amer. Psychologist*, 1948, **3**, 365. (Abstract.)
12. BLAKE, R., & MCCARTY, B. A comparative evaluation of the Bellevue-Wechsler Mental Deterioration Index distributions of Allen's brain injured patients and of normal subjects. *J. clin. Psychol.*, 1948, **4**, 415–418.
13. BOEHM, ALICE, & SARASON, S. Does Wechsler's formula distinguish intellectual deterioration from mental deficiency? *J. abnorm. soc. Psychol.*, 1947, **42**, 356–358.
14. BRECHER, SYLVIA. The value of diagnostic signs for schizophrenia on the Wechsler-Bellevue intelligence test. *Psychiat. Quart. Suppl.*, 1946, **20**, 58–64.
15. BROWN, A. W. Review of Rapaport, D. *et al.*, *Diagnostic psychological testing*. *Psychol. Bull.*, 1946, **43**, 477–479.
16. BROWN, M. N. A critique of the Wechsler-Bellevue system of weighted scores. *J. clin. Psychol.*, 1949, **5**, 170–173.
17. BUCK, J. N. *The H-T-P* (manual). Colony, Va.: Lynchburg State Colony, 1947. (Mimeo.)
18. CARP, A. Performance on the Wechsler-Bellevue Scale and insulin shock therapy. *J. abnorm. soc. Psychol.*, 1950, **45**, 127–136.
19. CHESROW, E. J., WASIKA, P. H., & REINITZ, A. H. A psychometric evaluation of aged white males. *Geriatrics*, 1949, **4**, 169–177.
20. CLARK, J. H. An investigation of certain relationships between the California Test of Mental Maturity and the Wechsler-Bellevue Intelligence Scale. *J. gen. Psychol.*, 1949, **41**, 21–25.
21. CLARK, J. H. Subtest variation on the Wechsler-Bellevue for two institutionalized behavior problem groups. *Amer. Psychologist*, 1949, **4**, 395. (Abstract.)
22. COFER, C. N. Psychological test performance under hyoscine: A case of post-infectious encephalopathy. *J. gen. Psychol.*, 1947, **36**, 221–228.
23. COHEN, B. Validity of a short form of the Wechsler-Bellevue on four psychiatric groups. *Amer. Psychologist*, 1946, **1**, 464. (Abstract.)
24. COHEN, B., & STEISEL, I. An evaluation of the various short forms of the Wechsler-Bellevue test. *Proc. Ia. Acad. Sci.*, 1947, **54**, 221–226.
25. COPPLE, G. Senescent decline on the Wechsler-Bellevue Intelligence Scale. Unpublished Doctor's dissertation, Univ. Pittsburgh, 1948.
26. COTZIN, M., & GALLAGHER, J. J. Validity of short forms of the Wechsler-Bellevue Scale for mental defectives. *J. consult. Psychol.*, 1949, **13**, 357–365.
27. CUMMINGS, S. B., JR., MACPHEE, H. M., & WRIGHT, H. F. A rapid method of estimating the IQ's of subnormal white adults. *J. Psychol.*, 1946, **21**, 81–89.
28. CUTTS, R. A., & SLOAN, W. Test patterns of adjusted defectives on the Wechsler-Bellevue test. *Amer. J. ment. Def.*, 1945, **50**, 98–101.
29. CUTTS, R. A., & O'KELLEY, MARGARET. The effect of hospitalization on Wechsler-Bellevue subtest scores by mental defectives. *Amer. J. ment. Def.*, 1947, **51**, 391–393.
30. DERNER, G. F., & ABORN, M. An administration board for the Wechsler-Bellevue object assembly subtest. *J. consult. Psychol.*, 1950, **14**, 71–72.

31. DIAMOND, S. The Wechsler-Bellevue Intelligence Scales and certain vocational aptitude tests. *J. Psychol.*, 1947, **24**, 279–282.
32. DUBIN, S. S., & THALER, M. The use of psychological tests on schizophrenic patients before and after shock treatment. *Amer. Psychologist*, 1947, **2**, 283. (Abstract.)
33. ESTES, S. G. Deviations of Wechsler-Bellevue subtest scores from vocabulary level in superior adults. *J. abnorm. soc. Psychol.*, 1946, **41**, 226–228.
34. FISHER, K. A. Changes in test performance of ambulatory depressed patients undergoing electric shock therapy. *J. gen. Psychol.*, 1949, **41**, 195–232.
35. FISHER, S., & SUNUKJIAN, HELEN. Intellectual disparities in a normal group and their relationship to disturbance. *J. clin. Psychol.*, 1950, **6**, 288–290.
36. FOSTER, A. Age and the Wechsler-Bellevue scattergraph. *J. clin. Psychol.*, 1947, **3**, 396–397.
37. FOX, CHARLOTTE, & BIRREN, J. E. Intellectual deterioration in the aged: Agreement between the Wechsler-Bellevue and the Babcock-Levy. *J. consult. Psychol.*, 1950, **14**, 305–310.
38. FRANKLIN, J. C. Discriminative value and patterns of the Wechsler-Bellevue Scales in the examination of delinquent Negro boys. *Educ. psychol. Measmt*, 1945, **5**, 71–85.
39. FREEMAN, A. V. Scatter indices in schizophrenia. Unpublished Doctor's dissertation, Columbia Univ., 1948.
40. FRENCH, ELIZABETH, & HUNT, W. A. The Navy Northwestern Successive Matrices Test. *Amer. Psychologist*, 1949, **4**, 268–269. (Abstract.)
41. GARFIELD, S. L. The Shipley-Hartford Retreat Scale as a quick measure of mental status. *J. consult. Psychol.*, 1947, **11**, 148–150.
42. GARFIELD, S. L. A note on Patterson's article, "The Wechsler-Bellevue Scale as an aid in psychiatric diagnosis." *J. clin. Psychol.*, 1947, **3**, 198–200.
43. GARFIELD, S. L. A preliminary appraisal of Wechsler-Bellevue scatter patterns in schizophrenia. *J. consult. Psychol.*, 1948, **12**, 32–36.
44. GARFIELD, S. L. An evaluation of Wechsler-Bellevue patterns in schizophrenia. *J. consult. Psychol.*, 1949, **13**, 279–287.
45. GARFIELD, S. L., & FEY, W. A comparison of the Wechsler-Bellevue and Shipley-Hartford scales as measures of mental impairment. *J. consult. Psychol.*, 1948, **12**, 259–264.
46. GARRETT, H. E. Review of Rapaport, D. *et al.*, *Diagnostic psychological testing*. *J. abnorm. soc. Psychol.*, 1947, **42**, 134–137.
47. GEIL, C. A. A clinically useful abbreviated Wechsler-Bellevue Scale. *J. Psychol.*, 1945, **20**, 101–108.
48. GERSTEIN, R. A. A suggested method for analyzing and extending the use of Bellevue-Wechsler vocabulary responses. *J. consult. Psychol.*, 1949, **13**, 366–370.
49. GIBBY, R. G. A preliminary survey of certain aspects of Form II of the Wechsler-Bellevue Scale as compared to Form I. *J. clin. Psychol.*, 1949, **5**, 165–169.
50. GILHOOLY, F. M. The relationship between variability and ability on the Wechsler-Bellevue. *J. consult. Psychol.*, 1950, **14**, 46–48.
51. GILHOOLY, F. M. Correction of "The relationship between variability and ability on the Wechsler-Bellevue." *J. consult. Psychol.*, 1950, **14**, 329.
52. GILHOOLY, F. M. Wechsler-Bellevue reliability and validity of certain diagnostic signs of the neuroses. *J. consult. Psychol.*, 1950, **14**, 83–87.

UNIVERSITY OF MICHIGAN LIBRARIES

53. GLASER, N. M. A study of the intelligence of immigrants. *Amer. Psychologist*, 1949, **4**, 241. (Abstract.)
54. GOLDFARB, W. Adolescent performance in the Wechsler-Bellevue Intelligence Scales and the Revised Stanford-Binet examination, Form L. *J. educ. Psychol.*, 1944, **35**, 503–507.
55. GOLDMAN, R., GREENBLATT, M., & COON, G. P. Use of the Wechsler-Bellevue Scale in clinical psychiatry with particular reference to cases with brain damage. *J. nerv. ment. Dis.*, 1946, **104**, 144–179.
56. GOTHBERG, LAURA. A comparative study of the Stanford-Binet old form test and Wechsler-Bellevue, verbal performance, and full-scale, as shown in the results of unselected employees. *Amer. J. ment. Def.*, 1949, **53**, 497–503.
57. GUERTIN, W. H. Mental growth in pseudo-feeblemindedness. *J. clin. Psychol.*, 1949, **5**, 414–418.
58. GURVITZ, M. S. An alternate short form of the Wechsler-Bellevue test. *Amer. J. Orthopsychiat.*, 1945, **15**, 727–733.
59. GUTMAN, BRIGETTE. The application of the Wechsler-Bellevue Scale in the diagnosis of organic brain disorders. *J. clin. Psychol.*, 1950, **6**, 195–198.
60. HALPERN, FLORENCE. Studies of compulsive drinkers: Psychological test results. *Quart. J. Stud. Alcohol*, 1946, **6**, 468–479.
61. HAMISTER, R. The test-retest reliability of the Wechsler-Bellevue intelligence test for a neuropsychiatric population. *J. consult. Psychol.*, 1949, **13**, 39–43.
62. HARPER, E. A. Discrimination of the types of schizophrenia by the Wechsler-Bellevue Scale. *J. consult. Psychol.*, 1950, **14**, 290–296.
63. HAYS, W. A comparison of scatter patterning for mental defectives on the Wechsler Forms I and II. *Amer. J. ment. Def.*, 1950, **55**, 264–268.
64. HERMANN, M., & HACKMAN, R. B. Distribution of scales on the Wechsler-Bellevue Scales and the California Test of Mental Maturity at a VA Guidance Center. *J. appl. Psychol.*, 1948, **32**, 642–648.
65. HEWSON, LOUISE. The Wechsler-Bellevue Scale and the Substitution Test as aids in neuro-psychiatric diagnosis. *J. nerv. ment. Dis.*, 1949, **109**, 158–183; 246–265.
66. HEYER, A. W. "Scatter analysis" techniques applied to anxiety neurotics from a restricted culture-educational environment. *J. gen. Psychol.*, 1949, **40**, 155–166.
67. HODGSON, G. An analysis of subtests in the Wechsler-Bellevue verbal scale administered to 139 delinquent Mexican boys. *Amer. Psychologist*, 1948, **3**, 343. (Abstract.)
68. HOLZBERG, J. D., & DEANE, M. A. The diagnostic significance of an objective measure of intratest scatter on the Wechsler-Bellevue Intelligence Scale. *J. consult. Psychol.*, 1950, **14**, 180–188.
69. HUNT, W. A., KLEBANOFF, S. G., MENSH, I. N., & WILLIAMS, M. The validity of some abbreviated individual intelligence scales. *J. consult. Psychol.*, 1948, **12**, 48–52.
70. HUNT, W. A., FRENCH, ELIZABETH, & MENSH, I. N. An abbreviated form of the McGill Verbal Situation Test. *J. clin. Psychol.*, 1948, **4**, 406–411.
71. HUNT, W. L. The relative rates of decline of Wechsler-Bellevue "Hold" and "Don't Hold" tests. *J. consult. Psychol.*, 1949, **13**, 440–443.
72. JASTAK, J. An item analysis of the Wechsler-Bellevue tests. *J. consult. Psychol.*, 1950, **14**, 88–94.
73. JOHNSON, ANNA P. Measuring men-

tal deterioration by the differential test score method. *J. ment. Def.*, 1947, **51**, 389–390.

74. JOHNSON, L. Wechsler-Bellevue pattern analysis in schizophrenia. *J. consult. Psychol.*, 1949, **13**, 32–33.
75. JOHNSON, T. F. Some needs in research with the Wechsler-Bellevue Scale. *J. gen. Psychol.*, 1949, **41**, 33–36.
76. KASS, W. Wechsler's Mental Deterioration Index in the diagnosis of organic brain disease. *Trans. Kans. Acad. Sci.*, 1949, **52**, 66–70.
77. KESSLER, LUCILLE. Intellectual changes in schizophrenic patients following electroshock therapy. Unpublished Master's thesis, New York Univ., 1947.
78. KLEIN, G. S. An application of the multiple regression principle to clinical prediction. *J. gen. Psychol.*, 1948, **38**, 159–179.
79. KLUGMAN, S. The effect of placement of the digits test in the Wechsler-Bellevue Intelligence Scale. *J. consult. Psychol.*, 1948, **12**, 345–348.
80. KNOTT, J. R., & CANNICOTT, R. The application of brief tests of intelligence in the psychiatric clinic. *Amer. Psychologist*, 1948, **3**, 365. (Abstract.)
81. KOGAN, KATE L. The diagnosis of a patient with organic defect. *J. Personality*, 1946, **15**, 113–120.
82. KOGAN, W. S. An investigation into the relationship between psychometric patterns and psychiatric diagnosis. *J. gen. Psychol.*, 1950, **43**, 17–46.
83. KRIEGMAN, G., & HANSEN, F. W. VIBS: A short form of the Wechsler-Bellevue Intelligence Scale. *J. clin. Psychol.*, 1947, **3**, 209–216.
84. KRUGMAN, M. Review of Rapaport, D., *et al.*, *Diagnostic psychological testing*. *J. consult. Psychol.*, 1946, **10**, 285–286.
85. KUTASH, S. B. A comparison of the Wechsler-Bellevue and Revised Stanford-Binet scale for adult defective delinquents. *Psychiat. Quart.*, 1945, **19**, 677–685.
86. LEVI, J., OPPENHEIM, S., & WECHSLER, D. Clinical use of the Mental Deterioration Index of the Bellevue-Wechsler Scale. *J. abnorm. soc. Psychol.*, 1945, **40**, 405–407.
87. LEVINE, J., & BLACKBURN, A. R. Intelligence test scores of newly blinded soldiers. *J. consult. Psychol.*, 1950, **14**, 311–315.
88. LEVINE, L. S. The utility of Wechsler's pattern in the diagnosis of schizophrenia. *J. consult. Psychol.*, 1949, **13**, 28–31.
89. LEWINSKI, R. J. Discriminative value of the sub-tests of the Bellevue verbal scale in the examination of naval recruits. *J. gen. Psychol.*, 1944, **31**, 95–99.
90. LEWINSKI, R. J. The psychometric pattern: I. Anxiety neurosis. *J. clin. Psychol.*, 1945, **1**, 214–221.
91. LEWINSKI, R. J. The psychometric pattern: II. Migraine. *Psychiat. Quart.*, 1945, **19**, 368–376.
92. LEWINSKI, R. J. The psychometric pattern: III. Epilepsy. *Amer. J. Orthopsychiat.*, 1947, **17**, 714–722.
93. LEWINSKI, R. J. Vocabulary and mental measurement: A quantitative investigation and review of research. *J. genet. Psychol.*, 1948, **72**, 247–281.
94. LINDNER, R. M., & GURVITZ, M. Restandardization of the Revised Beta Examination to yield the Wechsler type of I.Q. *J. appl. Psychol.*, 1946, **30**, 649–658.
95. LORGE, I. Intellectual changes during maturity and old age. *Rev. educ. Res.*, 1947, **17**, 326–332.
96. LUBORSKY, L. Psychometric changes during electric shock treatment. *J. nerv. ment. Dis.*, 1948, **107**, 531–536.

UNIVERSITY OF MICHIGAN LIBRARIES

97. LUCHINS, A. S. On certain misuses of the Wechsler-Bellevue Scales. *J. consult. Psychol.*, 1946, 10, 109–111.
98. MADONICK, M. J., & SOLOMON, M. The Wechsler-Bellevue Scale in individuals past sixty. *Geriatrics*, 1947, 2, 34–40.
99. MAGARET, ANN, & SIMPSON, MARY. A comparison of two measures of deterioration in psychotics. *J. consult. Psychol.*, 1948, 12, 265–270.
100. MAYMAN, M. An analysis of scatter in intelligence test results: A review of the literature. *Trans. Kans. Acad. Sci.*, 1946, 48, 429–444.
101. MCCULLOUGH, M. W. Wechsler-Bellevue changes following prefrontal lobotomy. *J. clin. Psychol.*, 1950, 6, 270–273.
102. MCNEMAR, Q. On abbreviated Wechsler-Bellevue scales. *J. consult. Psychol.*, 1950, 14, 79–81.
103. MCNEMAR, Q. Review of Rapaport, D., *et al.*, *Diagnostic psychological testing. Amer. J. Psychol.*, 1946, 59, 306–311.
104. MACPHEE, H. M., WRIGHT, H. F., & CUMMINGS, S. B., JR. The performance of mentally subnormal rural southern Negroes on the verbal scale of the Bellevue intelligence examination. *J. soc. Psychol.*, 1947, 25, 217–229.
105. OLCH, DORIS. Psychometric pattern of schizophrenics on the Wechsler-Bellevue intelligence test. *J. consult. Psychol.*, 1948, 12, 127–136.
106. PASCAL, G., & ZEAMAN, JEAN. A note on the validity of Wechsler-Bellevue scatter. *Amer. J. Psychiat.*, 1949, 105, 840–842.
107. PATTERSON, C. H. A comparison of various "short forms" of the Wechsler-Bellevue Scale. *J. consult. Psychol.*, 1946, 10, 260–267.
108. PATTERSON, C. H. The Wechsler-Bellevue Scale as an aid in psychiatric diagnosis. *J. clin. Psychol.*, 1946, 2, 348–353.
109. PATTERSON, C. H. A further study of two short forms of the Wechsler-Bellevue Scale. *J. consult. Psychol.*, 1948, 12, 147–152.
110. PEIXOTTO, HELEN. Wechsler-Bellevue patterns: A note of caution. *J. clin. Psychol.*, 1950, 6, 188–190.
111. RABIN, A. I. The use of the Wechsler-Bellevue Scales with normal and abnormal persons. *Psychol. Bull.*, 1945, 42, 410–422.
112. RABIN, A. I., DAVIS, J. C., & SANDERSON, M. H. Item difficulty on some Wechsler-Bellevue subtests. *J. appl. Psychol.*, 1946, 30, 493–500.
113. RABIN, A. I. Effects of electroshock treatment upon some aspects of personality and intellect. *Amer. Psychologist*, 1947, 2, 204. (Abstract.)
114. RAPAPORT, D., GILL, M., & SCHAFER, R. *Diagnostic psychological testing*, Vol. I. Chicago: Year Book Publishers, 1945.
115. RAPPAPORT, S. R., & WEBB, W. B. An attempt to study intellectual deterioration by premorbid and psychotic testing. *J. consult. Psychol.*, 1950, 14, 95–98.
116. RASHKIS, H. A., & WELSH, G. S. Detection of anxiety by use of the Wechsler Scale. *J. clin. Psychol.*, 1946, 2, 354–357.
117. ROGERS, L. S. A comparative evaluation of the Wechsler-Bellevue Mental Deterioration Index for various adult groups. *J. clin. Psychol.*, 1950, 6, 199–202.
118. ROGERS, L. S. A note on Allen's index of deterioration. *J. clin. Psychol.*, 1950, 6, 203.
119. RUBINSTEIN, E. A note on recording block design performance of the Wechsler-Bellevue Scales. *J. clin. Psychol.*, 1948, 4, 307–308.
120. SAPPENFIELD, B. R. Rapid method for placement of Wechsler object

assembly pieces. *J. clin. Psychol.*, 1947, **3**, 301.

121. SARTAIN, A. Q. A comparison of the New Revised Stanford-Binet, the Bellevue Scale, and certain group tests of intelligence. *J. soc. Psychol.*, 1946, **23**, 237–239.
122. SCHLOSSER, J. R., & KANTOR, R. E. Wechsler's deterioration ratio in psychoneurosis and schizophrenia. *J. consult. Psychol.*, 1949, **13**, 108–110.
123. SCHMIDT, D. G. Levels of intelligence of prison inmates. *Amer. J. ment. Def.*, 1946, **51**, 63–66.
124. SHOBEN, E. J. The Wechsler-Bellevue in the detection of anxiety: A test of the Rashkis-Welsh hypothesis. *J. consult. Psychol.*, 1950, **14**, 40–45.
125. SLOAN, W. Validity of Wechsler's deterioration quotient in high grade mental defectives. *J. clin. Psychol.*, 1947, **3**, 287–288.
126. SLOAN, W., & CUTTS, R. A. Test patterns of defective delinquents on the Wechsler-Bellevue test. *Amer. J. ment. Def.*, 1945, **50**, 95–97.
127. SLOAN, W., & GUERTIN, W. H. A comparison of H-T-P and Wechsler-Bellevue IQ's in mental defectives. *J. clin. Psychol.*, 1948, **4**, 424–426.
128. SORSBY, F. A briefer method for the Wechsler-Bellevue vocabulary subtest. Unpublished Master's thesis, Univ. Pittsburgh, 1949.
129. SPACHE, G. Scoring qualitative responses on the Wechsler-Bellevue Scale. *Amer. J. Orthopsychiat.*, 1948, **18**, 360–363.
130. SPRINGER, N. N. A short form of the Wechsler-Bellevue intelligence test as applied to Naval personnel. *Amer. J. Orthopsychiat.*, 1946, **16**, 341–344.
131. STACEY, C. L., & LEVIN, JANICE. Performance of retarded individuals on Stanford-Binet and Wechsler-Bellevue intelligence scales. *Amer. J. ment. Def.*, 1950, **55**, 123–132.
132. STROTHER, C. R. The performance of psychopaths on the Wechsler-Bellevue test. *Proc. Ia. Acad. Sci.*, 1944, **51**, 397–400.
133. TAGIURI, R. Comparison of results obtained from the Wechsler-Bellevue vocabulary test with those from the Stanford-Binet vocabulary test using a population of normal subjects and mental patients. *Bull. Canad. Psychol. Ass.*, 1946, **6**, 101. (Abstract.)
134. TEICHER, M. I., & SINGER, E. A report on the use of the Wechsler-Bellevue Scales in an overseas general hospital. *Amer. J. Psychiat.*, 1946, **103**, 91–93.
135. WARNER, S. J. The Wechsler-Bellevue psychometric pattern in anxiety neurosis. *J. consult. Psychol.*, 1950, **14**, 297–304.
136. WATSON, R. I. The use of the Wechsler-Bellevue Scales: A supplement. *Psychol. Bull.*, 1946, **43**, 61–68.
137. WEBB, W. B. A note on the Rabin ratio. *J. consult. Psychol.*, 1947, **11**, 107–108.
138. WECHSLER, D. *The measurement of adult intelligence.* (3rd Ed.) Baltimore: Williams and Wilkins, 1944.
139. WELSH, G. S. A note on scoring Wechsler-Bellevue subtests. *J. clin. Psychol.*, 1949, **5**, 421–422.
140. WHITEMAN, M. Altitude as a reference point in scatter analysis. *J. clin. Psychol.*, 1950, **6**, 160–164.
141. WILLIAMS, M., HUNT, W. A., & FRENCH, E. A further standardization of the CVS abbreviated intelligence scale. *Amer. Psychologist*, 1948, **3**, 365. (Abstract.)
142. WITTENBORN, J. R. An evaluation of the use of Bellevue-Wechsler subtest scores as an aid in psychiatric

diagnosis. *J. consult. Psychol.*, 1949, **13**, 433–439.

143. WRIGHT, H. F., MACPHEE, H. M., & CUMMINGS, S. B., JR. The relationship between the Kent EGY and the Bellevue verbal scales. *J. abnorm. soc. Psychol.*, 1949, **44**, 223–230.

144. WRIGHT, M. E. Use of the Shipley-Hartford Test in evaluating intellectual functioning of neuropsychiatric patients. *J. appl. Psychol.*, 1946, **30**, 45–50.

145. YARCOZYNSKI, G., BOSHES, B., & DAVIS, L. Psychological changes produced by frontal lobotomy. In J. F. Fulton *et al.*, *The frontal lobes*. Baltimore: Williams and Wilkins, 1948.

Received July 28, 1950.

A PSYCHOPHYSICAL METHOD EMPLOYING A MODIFICATION OF THE MÜLLER-URBAN WEIGHTS[1]

STEPHEN HARRISON
General Foods Corporation
AND
MARGARET J. HARRISON

The proposals made in this short paper have arisen out of taste and flavor threshold determinations carried out at these laboratories, but they should be applicable to psychophysics in general.

We have recently had the task of determining stimulus (absolute) thresholds for a number of different substances. Sometimes the threshold was measured for addition to spring water, but on some occasions it was necessary for our purposes to utilize an infusion of coffee as a base.

It has been the experience of this and many other laboratories working in the same field that organoleptic judgments are particularly prone to the "stimulus error"—i.e., the taster tends to align his judgment to his knowledge of the stimulus (2). It was, therefore, taken as an essential prerequisite of any procedure adopted that all samples should be coded before presentation to the subject. This restriction somewhat limited the choice of psychophysical methods, since it is a feature of much of the classical psychophysical procedure for absolute threshold determination that the subject has either knowledge or part knowledge of each stimulus presented to him. Stratton had proposed the method of serial groups, in which each stimulus value was presented together with an equal number of blank or "catch" stimuli, the task of the subject being to recognize the true stimuli. The mathematical treatment of the results leaves much to be desired, however, as the data are not fully utilized in deriving the threshold value and no parameter equivalent to a standard error is provided (7).

We have, therefore, developed the following method of estimation. In essence, coded beakers are presented in a series of pairs to the subject, each pair consisting of one beaker containing one of five concentration levels of the active substance under investigation and a second beaker serving as a "blank" stimulus containing water or other medium upon which the threshold determination is being based. No attempt will be made here to describe the details of the procedure, which is in

[1] The writers wish to acknowledge the cooperation of Dr. L. W. Elder, Director of the Produce Evaluation Laboratory of this organization, and of his test kitchen staff supervised by Mrs. A. M. Chao.

accordance with accepted psychophysical practice and follows recognized experimental design. The data are collected in the form of percentage of correct responses for each stimulus level presented. The mathematical treatment of the data, however, involves certain innovations in advance of standard methods which it is the purpose of this paper to present.

Method of Determining the Threshold

The method consists essentially in fitting a line to the data (after adjustment as described below), assuming the phi-log-gamma hypothesis (which is preferable to the phi-gamma hypothesis owing to the large value usually found for Weber's fraction for olfactory and gustatory modalities [8]). The technique of line-fitting used is that of least squares, as described, for example, by Guilford (6), but with three modifications.

1. A feature of the probit method as described by Bliss (3) and referred to by Finney (5) is incorporated into the analysis. This employs the least squares solution based upon an eye-fitted line. This has the particular advantage of permitting the use of zero and 100 per cent correct responses, which is not possible in the ordinary least squares treatment as described in the psychophysics literature.

2. A special allowance has to be made for the fact that a score of 50 per cent correct judgments would be expected by chance alone for each stimulus level. The original score should, therefore, be given a simple transformation so that it may express the percentage of the maximum possible score after chance expectancy has been discounted. This may be done as follows:

$$P = 2(p - 50)\%,$$

where P is the transformed percentage and p the original percentage of correct responses. The transformed percentages will then range from zero to 100 per cent (where zero indicates complete absence of discrimination and 100 per cent indicates perfect discrimination) provided the convention is adopted of interpreting negative transformed percentages as zero. This appears to the writers to be a reasonable convention since there would seem to be no way in which a taster could have negative discrimination. That is to say, it is assumed that original scores of less than 50 per cent (giving negative transformed percentages) can arise only from chance variations.

3. The Müller-Urban weights as normally employed in the constant process will clearly not be applicable here. The reason for this is as follows:

The Müller weight has the function of compensating for the fact that fluctuations of scores in the region of the threshold (i.e., the 50 per cent point) have less of a disturbing effect upon the location of the threshold value than high or low scores (4). Müller suggested that this should be compensated for by means of a weighting system based upon the squared ordinate of the normal curve. Hence, the Müller weights can properly be applied to the transformed percentage scores here proposed, since the transformed percentage scale ranges

from zero to 100 per cent with the threshold at 50 per cent (whereas the scale before transformation ranges from 50 per cent—effectively—to 100 per cent, with the threshold point at 75 per cent). The Urban weights correct for the fact that the error variance of a percentage score is not a constant but varies with the magnitude of the score, being greatest at the 50 per cent point and diminishing to zero as the score approaches zero or 100 per cent (**4**). Urban suggested that this should be compensated for by basing a weighting system on the reciprocal of the variance of a percentage, i.e., $1/pq$. This latter system of weights clearly cannot be applied *directly* to the transformed percentages since the error

TABLE I

MODIFIED MÜLLER-URBAN WEIGHTS

P %	*Weight*	*P %*	*Weight*	*P %*	*Weight*	*P %*	*Weight*
1	0.003	26	0.475	51	0.904	76	0.962
2	.010	27	.495	52	.915	77	.952
3	.019	28	.512	53	.925	78	.941
4	.031	29	.538	54	.935	79	.929
5	.045	30	.559	55	.944	80	.915
6	.060	31	.579	56	.953	81	.903
7	.076	32	.599	57	.960	82	.884
8	.094	33	.619	58	.968	83	.866
9	.112	34	.638	59	.974	84	.845
10	.131	35	.657	60	.980	85	.824
11	.151	36	.675	61	.986	86	.800
12	.171	37	.694	62	.990	87	.773
13	.191	38	.712	63	.994	88	.745
14	.213	39	.730	64	.996	89	.715
15	.234	40	.747	65	0.999	90	.681
16	.256	41	.763	66	1.000	91	.645
17	.277	42	.780	67	1.000	92	.605
18	.299	43	.796	68	1.000	93	.561
19	.321	44	.811	69	0.999	94	.512
20	.343	45	.826	70	.997	95	.458
21	.365	46	.840	71	.993	96	.399
22	.387	47	.854	72	.989	97	.329
23	.409	48	.867	73	.984	98	.249
24	.431	49	.880	74	.978	99	.151
25	0.453	50	0.892	75	0.970	100	0.000

$P = 2(p - 50)\%$, where p is the original percentage of correct responses.

variances upon which the Urban weights are based depend upon the *original* rather than the transformed percentages. Normally, the Müller and Urban weights are multiplied together to give a combined Müller-Urban weight, and it is in this form that they are usually tabled for use. These will not be applicable to the data under consideration since, as explained above, the Müller weight should properly be applied to the transformed percentage scores and the Urban

UNIVERSITY OF MICHIGAN LIBRARIES

weights to the original percentage scores. Instead, therefore, of using the normal Müller-Urban weights, it is suggested that a modified product of the two weights be used where the Müller weight is based upon the *transformed* percentage score and the Urban weight on the *original* percentage score. A table of these weights has been constructed for scores ranging from zero to 100 per cent at 1 per cent intervals (see Table I). To make these weights directly comparable to the ordinary Müller-Urban weights, they have beeen adjusted so as to bring the maximum weight up to unity.

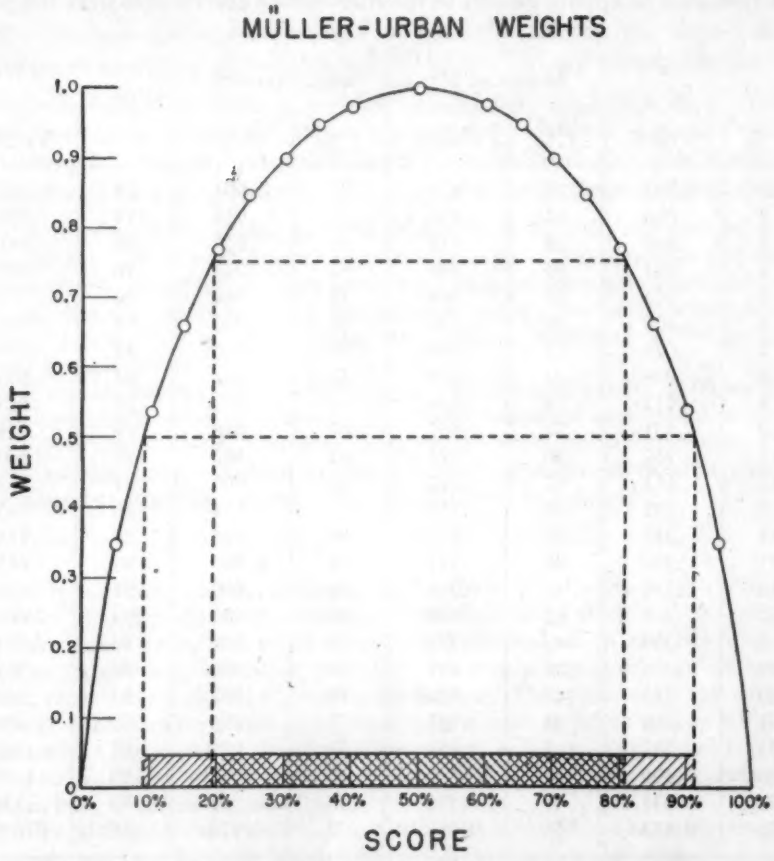

FIG. 1. CURVE REPRESENTING STANDARD MÜLLER-URBAN WEIGHTS

The standard procedure for the calculation of the threshold as given by Guilford (6) or by Bliss (3) has been followed except that calculations are based upon transformed percentages, and the modified Müller-Urban weights are used instead of the standard weights. These calculations

provide a measure of the standard error as well as the threshold, which makes it possible to establish fiducial limits for the threshold. This seems to the present writers to be a matter of great importance since the estimate of the threshold has little value by itself owing to the large

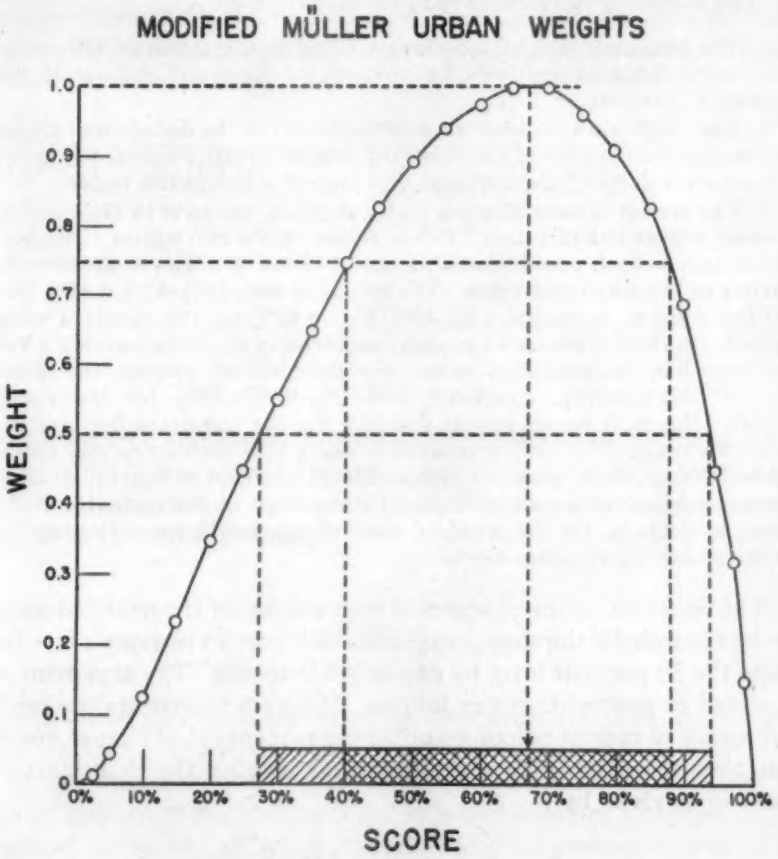

FIG. 2. CURVE REPRESENTING MODIFIED MÜLLER-URBAN WEIGHTS
The percentage score has been given transformation.

standard errors characteristic of thresholds in the fields of taste and flavor. A threshold determination is of little practical use unless it can be contained within fiducial limits. It has been chiefly this circumstance which has led the authors to develop this modified psychophysical procedure.

Comparison of the New Weights with the Normal Müller-Urban Weights

It is interesting to compare the properties of this new weighting distribution with the normal Müller-Urban weights (see Figures 1 and 2). The following differences may be noted:

1. The maximum weight is no longer found at the threshold (50 per cent) point, but is displaced positively, i.e., towards the upper end of the scale, falling at about 67 per cent.

2. The weights corresponding to the lower half of the distribution are much smaller than in the case of the standard Müller-Urban weights, owing to the high error variances of the corresponding original scores in this region.

3. The spread of scores having high weights is less than in the case of the standard Müller-Urban values. This is shown on the two figures. The heavily shaded area in each case includes all scores carrying a weight at least three-quarters of the maximum value. The spread is only 40%–87% = 47% for the modified weights, compared with 19%–81% = 62% for the standard weights. Similarly, the half-shaded area in each case includes all scores carrying a weight of at least half the maximum value. For the modified weights, the spread is only 27%–94% = 67%, compared with 9%–91% = 82% for the standard weights. Hence, it would appear desirable for the transformed scores to fall within the range 27%–94% if possible in order that each score may receive a high weighting factor. However, too restricted a range has been found undesirable since this gives a poor reliability for the slope of the regression line and hence, particularly, for the standard error of concentrations corresponding to the more extreme response scores.

The restricted range of scores of high weight for the modified weights can be expanded if the chance expectation of correct responses is reduced below the 50 per cent level by any suitable means. The argument may be stated in general terms as follows. Where p represents the original percentage of correct responses and e the percentage of chance expectation, then the percentage score after allowance for the chance expectation will be given by

$$P = (p - e)\left\{\frac{100}{100 - e}\right\}$$

so that

$$p = \frac{P(100 - e)}{100} + e.$$

The proposed weight is given by

$$\frac{y^2}{pq},$$

where y is the ordinate of the Gaussian curve of errors, the value of y being determined by the transformed percentage P; $q = 100 - p$.

Expressing the denominator in terms of the transformed percentage P, the expression becomes

$$\frac{y^2}{\left\{\frac{P(100-e)}{100}+e\right\}\left\{100-\frac{P(100-e)}{100}-e\right\}}$$

which may be written in the form

$$\frac{y^2}{\left(\frac{100-e}{100}\right)^2\left\{P+\frac{100e}{(100-e)}\right\}(100-P)}.$$

The squared expression in the denominator will be the same for all weights (for a given value of e) and so can be eliminated without affecting the relative values of the weights. The formula for the relative weights is, therefore, given by

$$\frac{y^2}{\left\{P+\frac{100e}{(100-e)}\right\}(100-P)}.$$

The numerator and denominator components of this weight now apply uniformly to the transformed percentage values.

As one means of reducing the chance expectation, the so-called triangular procedure might be utilized in which samples are presented in groups of three, two of which contain identical concentrations of the stimulus substance and the third, a blank solution only, the task of the subject being to identify the odd sample (1). Chance expectancy would then fall to $33\frac{1}{3}$ per cent, but the subject would have to examine one and one-half times as many samples as in the case of the method described above. In the field of organoleptics, onset of fatigue is particularly rapid, and for this reason we have not so far used the triangular test in systematic threshold determinations.

BIBLIOGRAPHY

1. BENGTSSON, K. Innebörden i statischt utslag vid provemakning. (Statistical evaluation of taste test results.) *Svensk bryggeritidskrift*, 1946, No. 7.
2. BENGTSSON, K., & HELM, E. Principles of taste testing. *Wallerstein Communications*, 1946, 9, 171–180.
3. BLISS, C. I. The calculation of the dosage-mortality curve. *Ann. appl. Biol.*, 1935, 22, 134–167.
4. BORING, E. G. Urban's tables and the method of constant stimuli. *Amer. J. Psychol.*, 1917, 28, 280–293.
5. FINNEY, D. J. The application of probit analysis to the results of mental tests. *Psychometrika*, 1944, 9, 31–39.
6. GUILFORD, J. P. *Psychometric methods*. New York: McGraw-Hill, 1936.
7. STRATTON, G. M. The method of serial groups. *Psychol. Rev.*, 1902, 9, 444–447.
8. THURSTONE, L. L. The phi-gamma hypothesis. *J. exp. Psychol.*, 1928, 11, 293–305.

Received July 6, 1950.

SOME COMMENTS ON THE CONTROLLING OF DIFFERENCES IN RESPONSES ON THE RORSCHACH TEST[1]

ROBERT M. EICHLER[2]
State University of Iowa

In a recent article in this journal, Cronbach (1) has called attention to the need for controlling individual differences in the number of responses given to the Rorschach test before proceeding to test the significance of relations involving the remaining Rorschach scores. As he indicates, the experimental dependence of the various Rorschach scores on "productivity" or number of responses (R) may result in a repetition of the test of the significance of relations involving differences in R when functions concerning other Rorschach categories are evaluated.

Since the typical method for controlling productivity differences by converting Rorschach scores to percentages presents serious statistical complications and, in addition, often fails to achieve the desired independence from R, Cronbach proposes several alternative procedures for resolving this difficulty. These are:

1. Scoring only a fixed number of responses on all protocols.
2. Reconstructing subgroups equated on R either by discarding cases or establishing subsamples for various intervals in which R is nearly uniform.
3. Plotting the variable against R and testing the significance of the proportion of cases falling above and below a line fitting the medians of the columns.
4. A method of "pattern tabulation" involving the analysis of profiles of normalized scores.

Cronbach's discussion of this particular problem and its possible solutions suggests two additional points that need consideration. The first relates to the inclusion of Rorschach scores in R (the problem of part-whole correlation); the second concerns the use of analysis of covariance with transformed Rorschach scores.

The dependence of a Rorschach score upon total responsiveness has usually been shown by correlating R and that score. Since the score is reflected in R, however, it is to be expected that the two will be positively correlated. It may be wondered, therefore, whether the Rorschach score is not actually independent of productivity when the latter does not include it. If this is true, the researcher would not be confronted with the task of controlling for R differences, before systematically

[1] The writer is indebted to Dr. Harold P. Bechtoldt for his helpful suggestions.

[2] Public Health Service Postdoctorate Research Fellow of the National Institute of Mental Health. Now at the New York Regional Office of the Veterans Administration.

UNIVERSITY OF MICHIGAN LIBRARIES

testing his hypotheses concerning the other test categories. Evidence obtained by the writer (2) indicates that, even with total productivity freed of a particular score, the correlation between the two tends to be quite high. For example, the study (60 cases) revealed correlations between weighted shading responses and *R minus* shading responses of .53, between *F* and *R minus F* of .49, between weighted color responses and *R minus* color responses of .38, between *P* and *R minus P* of .37. Thus, even when one avoids any spurious correlation, a "fluency" factor is apparent. In view of these significant relationships it is manifestly an important first step, as Cronbach suggests, to control for productivity differences if obtained differences in other Rorschach categories are to be interpreted soundly.

In his discussion of procedures for achieving independence from *R*, Cronbach has apparently not considered the possibility of using analysis of covariance. The techniques he suggests require either a large number of cases or a considerable number of responses per blot if they are to be useful. Since the administration and scoring of a Rorschach test demand a good deal of time, the usual researcher in this area generally uses a sample fairly limited in size, with the resultant *N* for subgroups even more limited. For such investigations covariance analysis would seem to be better adapted to handle the problem of productivity control. Such analysis makes the groups comparable on the productivity factor by determining how much of the variance in the particular Rorschach score can be predicted from total responsiveness, and then subtracting this to secure the residual variance as the adjusted value. It should be noted that if this method of analysis is employed, it will be generally necessary to normalize the data by some appropriate transformation, as most Rorschach scores are considerably skewed. The writer has found the square root transformation (4) a useful method in this regard. To avoid spurious correlations between productivity and other Rorschach scores, it is also suggested that the total number of responses *minus* the number of responses in the dependent variable (the Rorschach score under analysis) be used as the independent or uncontrolled variable.

Some qualifications concerning the use of analysis of covariance should be made. In a simple covariance design where the total variance is analyzed into the methods and within groups components, the extent to which this technique will increase the precision of an experiment will depend upon the within groups correlation between productivity and the Rorschach score. Where this correlation is low and/or where the group differences on the unadjusted Rorschach score under analysis are small

and attributable to chance, the use of this design may be of little utility. In cases where a more complex covariance design is used involving the interaction variance as the error term, the increase in precision obtained will depend upon the correlation of the interaction effects (3). Such a design will not be profitable where the latter correlation is low.

BIBLIOGRAPHY

1. CRONBACH, L. J. Statistical methods applied to Rorschach scores: A review. *Psychol. Bull.*, 1949, 46, 393–431.
2. EICHLER, R. M. Experimental stress and alleged Rorschach indices of anxiety. *J. abnorm. soc. Psychol.*, 1951, 46. (In press.)
3. LINDQUIST, E. F. *Statistical analysis in educational research.* Boston: Houghton Mifflin, 1940.
4. SNEDECOR, G. W. *Statistical methods.* Ames, Iowa: Iowa State Coll. Press, 1946.

Received July 5, 1950.

BOOK REVIEWS

BORING, E. G. *A history of experimental psychology.* (2nd Ed.) New York: Appleton-Century-Crofts, 1950. Pp. xxi+777. $6.00.

The second edition of his *History* is, in Boring's own words, "about one-third larger than the first edition, is one-half new writing and uses for its other half about two-thirds of the old edition. There must be some thousand minor changes in what is reprinted. . . . [Despite the changes] the old rule for the first edition and for my *Sensation and Perception in the History of Experimental Psychology* still holds: I speak with confidence up to twenty years ago (1930); I speak, but with less assurance, of the next decade; whatever I say for the most recent decade is based on gratuitous courage."

Changes have indeed been made. Chapter 11, in the early part of the book, is new and presents material on the Scottish school, French empiricism, and French materialism. The material of chapter 13 is extended by a consideration of Kant; and Hering receives amplified treatment in chapter 17.

The greatest changes occur in the last part. Chapters 22 through 26, dealing, respectively, with American functional psychology, Gestalt psychology, behavioristics, brain function, and dynamic psychology, have been completely rewritten.

Chapter 27, entitled "Retrospect," gives a valuable summary of important men and trends in the history of psychology. In assessing the great contributions, Boring says: "The student of the history of psychology would not be promoting an absurdity if he placed on the horizon of his imagination these three landmarks: Fechner's *Elemente der Psychophysik* of 1860, Ebbinghaus' *Über das Gedächtnis* of 1885, and Freud's *Die Traumdeutung* of 1900." (Many psychologists will want to know why Pavlov's *Conditioned Reflexes* is not included.)

In his 1929 edition Boring failed to find any great psychologists; now he says that there were four: Darwin, Helmholtz, James, and Freud. [I hereby nominate Pavlov!] In 1950, however, the *great-man* theory of scientific progress is abandoned; the great man is now regarded as only one of the agents of progress.

Although he states that he cannot speak confidently of the years since 1930, Boring gives us (in his Preface and in individual chapters, but especially chap. 27) an account of psychology in 1950. These are the striking elements: All American psychology is functional. Gestalt psychology has accomplished its mission and is now gone. Behaviorism and operationism have placed a stamp of objectivity upon American psychology. [It is disappointing not to find an account of H. M. Johnson's pioneer operational analyses (*Psychol. Rev.*, 1928, 1929, 1930).] Studies of brain function still go on [but little is said of their present significance]. Dynamic psychology is an interest, a concern, and a movement for "those psychologists who, disappointed in the description of consciousness and behavior, seek a more satisfying psychology of what they sometimes call 'human nature' "; it appeals to many psychologists; hence it constitutes a strong movement.

In 1950, psychology need make no apology. Through the years psychology

has gained self-confidence from the successful application of its facts and principles. All areas of psychology show encouraging growth. "In both Great Britain and America psychology's battle has been to become independent of philosophy. . . . That conflict between parent fixation and the need for independence was still felt in 1929, but now it has all gone." [But think of the number of courses in scientific method that are given in departments of psychology! Think also of the annual crop of "systematic" articles!]

One of the first things that impresses a reviewer in comparing the first and second editions of the *History* is the similarity of historical method used in the two volumes. The *great-man* theory is disavowed in 1950, but the historical analysis exemplified in the second edition is not greatly changed from that of the 1929 edition, where, presumably, the theory helped to determine the treatment.

It may have been the *great-man* view that produced, in the 1929 edition, a history emphasizing men, their controversies, and their secondary professional activities. The weighting of emphasis led Woodworth (*J. genet. Psychol.*, 1930, **38**, 521–526), in his review of the first edition, to say:

> The author's story of the past seventy years centers about the systematic problems and difficulties. He has much to say of the busy externals of psychology, the founding of laboratories, journals, and societies in the different countries, and the succession of important systematic books that have appeared. He is particularly full on the biographical side. To quite an extent, he presents the broadening of the scope of psychology as a consequence of the new methods invented. But he says little of results. He does not show how the psychologists increased the body of knowledge inherited from the sense physiologists, nor what new knowledge they have gained with their own new methods.

The situation is not greatly changed in the second edition. The *great-man* theory, although disavowed, has left its imprint.

Fortunately, we do not now have to share in Woodworth's frustration of 1930. Other works (for example, Murphy's *Historical Introduction to Modern Psychology*, are available to fill out the historical picture, and, in particular, we have Boring's *Sensation and Perception in the History of Experimental Psychology*, an account that treats experimental psychology as a factual science. The *History* and *Sensation and Perception* are complementary; and if a third book on learning and motivation should one day make its appearance (a possibility that Boring mentioned in his Preface to *Sensation and Perception*) we shall have a trilogy that gives an admirable account of a large segment of experimental psychology, written in a temporal context that varies, depending on the subject matter, from 1930 to the present.

Boring's insistence that he cannot speak confidently of the years since 1930 poses a difficulty for a reviewer: Should the reviewer ignore all that Boring says concerning the last twenty years, or should he give it serious attention? The appropriate answer seems to be that the discussion

is on the record; hence it cannot be ignored. A more indirect answer argues for its consideration because it provides a clue to the vantage point from which history was surveyed; hence it presents a background against which Boring's historical selections and emphases may be evaluated.

A history written in the context of 1930 may be expected to be different from one written in a context of 1950. Further, it may be expected that if the historian shifts his vantage point from time to time and from topic to topic within the twenty-year span, the shifting point of reference will give rise to different interpretations of historical trends.

Boring's tendency to shift his point of survey within the twenty-year span since 1930 is sometimes disconcerting. It is true that, in some instances, a more sympathetic account of a topic is achieved by considering its history in the context of 1930. (For example, it is likely that the contributions of the introspective psychologists will be regarded much more sympathetically in the earlier context than in the later. It is hard to imagine that the controversies which concerned these men hold great interest for us now, nor are we likely to get excited about the philosophical backgrounds of many of the controversies.) But in other instances, as, for example, the discussion of vision in both the *History* and *Sensation and Perception*, one would like a discussion of more recent material. (Hecht, whose work far outweighs in importance the contributions of most of the earlier workers, is not mentioned.) Without the discussion of recent results, the significance of the account is diminished. In any case, the temporal unevenness of the topical casting of accounts leaves us with a feeling that a considerable reworking of history since 1930 will be required before we have a rounded picture of the events of the last two decades and, possibly more important, before we can begin to visualize the effects of present points of reference upon the interpretation of earlier psychology.

However, a few generalizations seem to be in order. The years near 1930 were important years in the history of psychology. Psychology at that time was beginning to feel some influences that had remained small or latent, and the decade ending with 1930 marked the rapid recession from the high water mark of introspective psychology. Experimental psychology before 1930 was introspective; since 1930 it has been objective.

In establishing the objective nature of psychology, one influence is of great importance, the biological. Boring does not give a rounded account of this influence. Names are listed: Helmholtz, Pavlov, Loeb, Sherrington, Jennings, Adrian, and those of many other men; but the influences of the men named are not as clearly traced as are the influences of the psychologists. Psychologists since 1930 may deal in molar concepts, but in one main regard and in many minor regards, their approach is biological: they deal with behavior and not with mind. In the

study of behavior, many basic concepts and the emphasis on biological, objective method come from our biological rather than our psychological ancestors. Even the work on so-called sensation and perception, which might be expected to trace most readily back to psychological influences, is probably more influenced by the work of the sensory physiologists than it is by the work of sensory psychologists. And so in all areas of experimental psychology a primary legacy of methodology and concept is derived from a line of biological workers, many of them outside the general stream of early psychological history. In some future history it may be as necessary to trace those biological trends that helped determine American objective psychology as it is to follow the long path of introspective psychology. It may turn out that this type of history will emphasize such things as the way experiments were done and the way questions were asked rather than verbalized tenets of faith.

Another important influence that may be traced in a later history is the influence of psychiatry and the social sciences on experimental psychology. This influence may, in this volume, be pieced together on the basis of the discussion of many individuals and the chapter on dynamic psychology, but many psychologists will probably feel that nothing short of an account fully commensurate with that given introspective psychology will suffice. There can be little doubt that medicine and social science have had some influence on experimental psychology, and insofar as the general body of psychologists is concerned, one needs only to look at interests represented in the American Psychological Association to be convinced that, as of 1950, G. Stanley Hall has carried the day.

One cannot but be impressed anew with Boring's scholarship. It is probably true that no other psychologist could provide the high level of historical narration that Boring demonstrates in his *History*. The presentation is lively, the attention to detail meticulous, and the array of facts overwhelming. As Elliott says in the editor's introduction: "Even if psychologists are a crotchety lot, no audacious critic among them is going to rise and deny that Boring's *History* is a classic." The *History* is indeed a classic; but its outstanding quality derives more from an unparalleled treatment of the rise and fall of introspective psychology than it does from a delineation of the roots of objective psychology.

C. H. GRAHAM.

Columbia University.

ENGLE, T. L. *Psychology: Its principles and applications.* (Rev. Ed.) New York: World Book Co., 1950. Pp. xi+628. $3.08.

Authors of texts in elementary psychology have generally ignored the increasing interest in this subject as part of the high school curriculum. All too frequently general educational values inherent in the study of general psychology have been sacrificed in an attempt to provide an

introduction to knowledge and skills thought to be the necessary equipment of the bright college sophomore who might happily decide to complete an undergraduate major and eventually to acquire the Ph.D. degree in psychology.

This book differs from the usual introductory text in that it has been written for general students in both the final years of high school and the beginning years of college, who in most instances will have no further courses in the subject. It is a revision of the first edition (1945) and incorporates numerous modifications based on suggestions of teachers, questions and comments raised by the author's students, and the results of psychological research performed during and following World War II.

The 17 chapters are organized into six major units: "I. What is the Science of Psychology," "II. Patterns of Human Behavior," "III. Why We Have Our Varied Patterns of Behavior," "IV. Learning," "V. Problems of Personal Adjustment," "VI. Broader Problems of Adjustment—Society, Job, and Family."

The numerous contemporary illustrations and practical applications of psychological principles are likely to be interesting and useful for high school students. Despite the practical orientation, the material presented is drawn largely from experimental sources. The interpretation of this material is objective but eclectic. References to sources are omitted since "for the students for whom this book is intended such source material would have slight practical value, if any."

As a text for use on the college level, even if the general course is conceived in terms of its contribution to general education, this book seems to be somewhat superficial. As a high school text, however, it should prove to be highly satisfactory since it is written on an elementary and practical level and, at the same time, maintains an admirable degree of objectivity and scientific integrity.

FREDERICK A. COURTS.

Reed College.

GUTHRIE, EDWIN R., AND POWERS, FRANCIS F. *Educational psychology.* New York: Ronald Press, 1950. Pp. vi+530. $4.00.

This volume is designed as an introductory text in educational psychology addressed to prospective teachers. It stresses the psychology of learning and its applications to classroom problems.

Its special quality arises from the fact that is is dominated by Guthrie's theory of learning by association. Stimuli are goads to trial-and-error activity. Eventually the cat or the child finds the response which permits escape from an unpleasant situation. This correct response gets learned because it is the last of the series of responses and thus is protected from unlearning.

When a child enters school he confronts an enormous number of new skills which must be learned by trial and error. . . . He must learn how to sit rela-

tively still, how to remain relatively quiet. . . . In a large number of these skills, the motivating agent is the teacher. She is the goad that keeps prodding the child into activity until he achieves the required result (p. 120).

Guthrie's theory of learning has the respectability of being one of several recognized theories. The exposition of other theories is rather good but decidedly condescending. The implications of this theory for classroom teaching, as indicated by the above quotation, seem to me to be unfortunate.

Only rarely does another point of view emerge. Chapter XII begins with these lines: "The child is a whole and entire organism. He exists in a world of situations which are complex stimulus patterns. To these situations he responds as an entire organism." Chapter XIII begins, "A modern program of education calls for rich and varied learning experiences for every child." The promise of these words is not fulfilled.

There is much solid material in this text. For the most part the exposition is very clear and well organized. There are many homely references to actual classroom problems. There is a wealth of detail and of practical, common-sense advice.

The authors assert (p. 112) that different learning theories amount to much the same thing from the point of view of practical educational psychology. Their text is a demonstration to the contrary. The prospective teacher who absorbs this volume will be a different and, I fear, a poorer teacher of children.

FRANK K. SHUTTLEWORTH.

The City College of New York.

ANASTASI, ANNE, AND FOLEY, JOHN P., JR. *Differential psychology.* (Rev. Ed.) New York: Macmillan, 1949. Pp. xv+894. $5.50.

This revised edition has been greatly enlarged and improved over the original text. Extensive material from biology, child development, and anthropology has been added, as well as many new citations from the psychological literature, which has grown so extensively during the 12 years between the editions. While the general outline of the book remains the same, there is marked improvement in the organization of many chapters, in the use of graphs and tables, and in the general format.

In their preface, the authors propose to place major emphasis on methodological problems and basic concepts. This they have done. However, for a subject so involved with psychological theory, this text fails to present a clear-cut statement of principles and postulates which the authors obviously have assumed before undertaking their discussion of the literature. They attempt such a statement in the first four chapters but, in this reviewer's opinion, have failed to be consist-

ent. After discussing and seeming to accept the "interaction viewpoint" concerning heredity and environment, the authors give little recognition to this interdependence. They carry forward habits of thinking which reflect behaviorism much more clearly than the "field theory" which is often implied by the interaction viewpoint. Throughout, the authors' sympathies are clearly with environmental determinism.

The authors' inconsistency in criticism appears most clearly in discussion of studies establishing differences between groups. The authors seem eager to discount studies which have established racial, national group, or sex differences unless the demonstrated superiority of one group in measured abilities can be associated with obvious environmental advantages, as in Klineberg's studies. With such studies, as with studies which report no differences, the text finds little fault. In other studies, the sampling is usually attacked, or if intrinsic or native differences are implied, alternative explanations in terms of learning are advanced. (One possible exception may occur in the fact that while there are lengthy discussions of how the feebleminded child's low IQ might result from environmental influences and hence be a social product, the child prodigy is accepted as a sort of psychological curiosity, admittedly valuable, without any searching for the social or environmental factors which produced him!)

Studies are criticized in detail which attempt to allow for environmental advantages of one group and yet do find group differences (e.g., Tanser, Bruce). When behavior changes occur in the *absence* of special environmental advantages, as specifically the changes which occur in loss of skill in a practiced group during an interval of no practice, then the authors feel it necessary to advance several hypotheses as to why this loss could occur. (Obviously it shouldn't!) Where studies frankly avow a cultural origin of the differences observed, such as culturally-determined differences in expressive movements, the studies are accepted with little criticism of methodology or with no mention of alternative explanations. One of the few exceptions to this is found on page 652 in a reference to girls' "demonstrated superiority in linguistic aptitude," which is accepted without particular debate.

The authors reject "growth" as having much to do with the intellectual and emotional development of the older child, though they admit that growth may play a major role in the behavioral development of the very young child (p. 265 ff.). They do not cite, however, the difficulty of an individual's later escaping from the impact of this early growth once it has set or sensitized the learner toward certain classes of stimuli. At one point the authors hold that a psychological growth curve is "at best a practice curve obtained in the absence of controlled conditions" (p. 278). The authors then assert that such curves can vary in different cultures. This reviewer gained the impression that the

authors believe there is nothing particularly definitive or deterministic about the learning process, since learning curves are so subject to cultural vicissitudes. Yet an individual is the creature of his developmental *or* of his learning processes, and the cumulative significance of either process for the individual cannot be minimized.

The book is very complete in literature cited and in the coverage of its topic. While it is a little ponderous for undergraduates in psychology or education, it certainly is well suited to the advanced critical student. It deserves wide usage. Possibly the inconsistencies in criticism and interpretation which occur add to its merit for use with students who have some background in literature and in research method. It becomes a provocative source of discussion, where a more consistent book would be only doctrinaire.

DALE B. HARRIS.
Institute of Child Welfare, University of Minnesota.

HOPPOCK, ROBERT. *Group guidance: Principles, techniques, and evaluation.* New York: McGraw-Hill, 1950. Pp. xiv+393. $3.75.

This book is a practical handbook for guidance workers who plan to include group guidance procedures as a part of their total guidance program. The author avoids theoretical issues and makes the book consist primarily of clear, concise, practical suggestions regarding the content, techniques, and evaluation of group guidance. The content of the book is limited to the use of group guidance in orientation, educational guidance, and vocational guidance; other areas in which group guidance methods could be used are deliberately omitted. He makes it clear that the book is intended for teachers and school or college guidance workers but not intended for clinical psychologists and personal adjustment counselors. However, many of the techniques of group guidance and of evaluation of guidance could be utilized in group psychotherapy.

The book has four major parts: principles, techniques, evaluation, and appendixes. Part I, "Principles," discusses such practical questions as: Who should get it? Should it be required? Who should supervise it? What should it include? These questions are discussed only from the point of view of the use of group guidance within a school program.

Part II, "Techniques," outlines the purpose, preparation, and procedures of more than twenty techniques of group guidance, including group visits, student surveys, case conferences, visual aids, panel discussions, dramatizations, and others. The various techniques are concisely outlined and illustrated by summaries of articles from professional journals and other sources.

Part III, "Evaluation," stresses the need for objective evaluation of the outcomes of guidance. It consists primarily of brief summaries

of twenty research studies, thus presenting specific examples of evaluative techniques, the variety of criteria utilized in evaluation, and the many problems involved in attempting to evaluate guidance.

Part IV, "Appendixes," is the longest of the four parts. It contains primarily verbatim records of demonstrations of many of the group guidance techniques which are outlined in Part II and helps to clarify the points stated in the text.

The author is to be commended for his frankness, sincerity, and modesty. For example, in contrast with some other writers of guidance books, he states clearly that he does not consider group guidance as *the* guidance technique but regards it as effective only when coordinated with other guidance techniques. Also, he states that "it would be odd if future research did not reveal some of his methods and conclusions to be erroneous." However, persons who plan to use or who already are using group techniques will find the book a useful and practical guide for organizing their programs, selecting and applying their techniques, and for evaluating their work objectively.

WILBUR S. GREGORY.

University of Redlands.

PLANT, JAMES S. *The envelope: A study of the impact of the world upon the child.* New York: Commonwealth Fund, 1950. Pp. xi+299. $3.00.

It is a bit odd to read this last work by the author of *Personality and the Cultural Pattern.* The earlier work was a major attempt at synthesis and was the first clear expression of a movement which today incorporates psychologists, sociologists, anthropologists, and psychiatrists. The movement is known by the convenient, if ugly, expression, "Culture and Personality." The present work has none of the scope and depth of the theoretical work which it follows by some 13 years. Indeed it seems miscellaneous. But it only *seems* so. It is written around 21 problems, each of which takes a separate chapter. Illustrations of these problems are: security, authority, feeling of difference, intelligence, reaction to failure, minority race problems, narcissism, and the meaning of meaning. Each topic is discussed as an entity, but to it is applied very effectively the earlier-developed notions. The outcome is stimulating. The rich experience and insight of the clinician has been joined to the tolerance and humility of a person who sees and understands the complexity of society. There is no glibness, no false hopes; but there is no annoyance, either, at the reality picture of a dependent organism trying to adjust to shifting, inordinate demands and requirements. For him man is a temporary solution confronted by problems and challenges to these earlier adjustments.

The major theme, the "Envelope," refers to a large, integrated area of personality which interprets the social and physical world to the

organism and *through* which he communicates himself to others. The Envelope is comparable to such terms as autism and motivated perception, though it is not, of course, identical with them. All such terms stress that there are various selective processes at work which change, distort, and vary the "input" to make it congruent with the person's needs and, therefore, to protect him. Where Plant's special sensitivity (and sensibility) comes in is in his handling of the interrelationships between various needs. He sees them as changing, as sometimes not what they seem to be, as arising, functioning, and disappearing in a matrix of social and "accidental" variables. And by "needs" he does not mean basic physiological processes alone; he refers to the requirements which one finds in the clinical examination of problem children. It is because of this that he can think of relating to personality problems which arise in school situations the "problems arising out of the fact that steam cannot be efficiently transported . . . that people have to move to this source of power. Electricity and chemical sources, on the other hand, are decentralizing . . . and their development will mean other shifts in social pressures on the individual." In his hands such observations make clearer the job of the clinician and the difficulties he will face.

The only weakness in the volume is the lack of concern with learning as a process. Given motivation and a problem, the behavior equipment seems always to be available and ready. But within the context of clinical problems perhaps the weakness is not Plant's but of those to whom he would look for aid in their solution. In any event it is the kind of book which one reads, enjoys, and learns from.

RICHARD A. LITTMAN.

University of Oregon.

EDWARDS, ALLEN L. *Experimental design in psychological research.* New York: Rinehart, 1950. Pp. xiv+446. $5.00.

This is an excellent book which will be of great value as a text and reference. It does not, however, have the content which some readers would expect from the title. Its content is better indicated by the following statement from the preface:

> This book attempts to present to the student in psychology, education, sociology, and the behavioral sciences, some of the newer developments in statistical analysis, particularly with respect to small-sample theory, as they relate to problems of research and experimentation in these fields.

The objective expressed in this statement is attained in admirable fashion. The author presents a variety of methods of statistical analysis. They are primarily the ones that the experimenting psychologist today is apt to find most frequently useful, and include some that are not widely known to those who would find them useful. He organizes his

presentation primarily around types of research design, discussing in each case the analyses appropriate to the design. The presentation is admirably clear. In my opinion, this book would provide for very able students a first-rate introductory textbook on statistics (though, since it is obviously not intended for introductory use, some supplementation would be needed). Where the student cannot be given the advantage of such an introduction, the book should be highly useful as a text for an advanced course in statistics. Where a combined course is given on statistics and experimental method, this book provides a practically ideal text if the students are up to it. For advanced students and psychologists the book provides probably the best available general guide to the statistical analysis and evaluation of experimental data.

This book, despite its excellence, does not fully meet the need for a textbook on experimental design in psychology. To state this is not to criticize the author, since he apparently did not have this as his *main* aim in writing the book. In view of the title which the book bears, however, it seems important to state why the book is not fully adequate in relation to this need. Some of the reasons follow:

1. From the point of view of this need, the great emphasis on statistical analysis is distracting. For the student whose principal aim is to learn about research design, matters of statistical analysis should be treated in proper perspective as a tool, rather than be the main subject of discussion, if he is not to be diverted from his aim.

2. From the same point of view, there are some important omissions. The most glaring one is the absence of explicit general discussion of the frequently used design in which a subject is used as his own control and there is hence need to counterbalance experimental treatments and possible temporal effects. In the chapter on "Experimental Designs Involving Repeated Measurements of the Same Subjects," this design is not even mentioned. It is discussed only in the chapter on the Latin square, and there primarily as illustrative of the use of the Latin square rather than as a design involving some distinct problems of its own.

3. On problems of design in contrast with analysis, the author does not write at much length or with as great understanding and awareness of problems.

4. It appears to be something of a defect that the author has not introduced much of the terminology which has been developed by Fisher and others for aspects of research design. Psychologists and their students will presumably need to continue to consult the writings of these men. It would facilitate this reference if terms such as "randomized block" and "split plot" design had already been encountered in a book of this sort. These aspects of design are treated here, and no other single labels are offered for them as more appropriate in psychology. To have provided any label would have facilitated the student's getting a useful concept; to have provided the Fisherian label would have the added advantage of increasing the student's scientific literacy.

Despite these objections, the book represents a definite gain to the psychological literature, even on the subject of experimental design. It

doesn't do as much as one could hope, but it does something very valuable which has not been done before.

I have a few comments on specific parts of the book, which may be of use to some of its readers. They follow:

1. The author does a service in bringing into the psychological literature a useful convention of referring to one-tailed tests by the term *significance point* and two-tailed tests by the term *significance level*. The clarity and consistency with which this distinction is used here should help in eliminating a source of unnecessary confusion. It should be noted, however, that the distinction as used here may lead the student into fresh confusion at one point—the interpretation of F when used to test the significance of the difference between two means. The usual use of the F tables for this problem yields a significance *point* with respect to the distribution of F, but this is equivalent to a significance *level* with respect to the difference between the means (since a high value of F may be produced by a difference in either direction between the two means). This fact is not taken account of by the author, and the student who adopts Edwards' terminology will be puzzled when he finds that we do not then, as Edwards claims, "arrive at exactly the same conclusion that we reached by means of the t test" (p. 184). We would, in fact, arrive at conclusions which differed by a factor of 2. The confusion could be cleared up by applying the term *significance level* to the one-tailed F test in this particular case; if one wants for this case a *significance point* which will correspond to that obtained through t, it will be necessary to divide the tabled p value by 2.

2. The author's italicized statement on pages 216–217 is too sweeping and might well lead a student to make an incorrect test of significance in analyzing some of the designs discussed in chapters 15 and 16.

3. There seems to be a slightly misplaced emphasis in the author's suggestion (as on p. 5) that it is important to try to get *continuous* scales of measurement. Surely for almost all purposes we would gladly settle for *discrete* scales; there are other scale properties much more important to worry about than this one.

4. The author's statement of the meaning of triple interactions (pp. 227–228) is not at all clear. A person who does not already know their meaning is likely to understand from it that they simply confirm the existence of several first-order interactions.

5. The author's statement on page 257 that "We would thus be led to the false conclusion, in this analysis, that none of the main variables is significant," is misleading. If the hypothetical experiment referred to had been done in the way described here, this would be the true and appropriate conclusion.

6. It is unfortunate that Edwards' discussion of systematic Latin squares (pp. 307–310) is so inconclusive. He gives the impression that methodological research is the only source of usable knowledge about the bias of systematic squares, whereas purely formal knowledge might well be of value.

7. The value of the book is increased by the great diversity of examples described, both in the main text and in the problem section at the end of each chapter, by the numbered list of formulas printed just before the Appendix, and by the well-selected tables contained in the Appendix.

IRVIN L. CHILD

Yale University.

COCHRAN, WILLIAM G., & COX, GERTRUDE M. *Experimental designs.* New York: Wiley, 1950. Pp. ix+454. $5.75.

The authors attempt to ". . . describe in some detail the most useful of the designs which have been developed, with accompanying plans and an account of the experimental situations for which each design is most suitable" (p. v). They succeed admirably in achieving this and other objectives. It is assumed that the reader has had more than a casual exposure to analysis of variance, upon which virtually all of the analyses depend. There is no emphasis on non-parametric situations and specialized tests.

The first three chapters deal with: (1) the contribution of statistics to experimentation, and initial steps in planning experiments; (2) methods of increasing the accuracy of experiments, including replication, refinement of measures, design, and grouping data; and (3) notes on analysis of variance. The first two chapters contain much wisdom, stated with such simple humility that the presentation attains a homely beauty. In chapter III easy approaches are made to difficult topics. Likelihood ratio tests are gently introduced. Mundane but difficult matters such as accuracy of computations are treated with care born of long experience with real data.

Chapter IV deals with randomized designs, single, double, and triple groupings, and chapter V covers factorial experiments. Each design is discussed under: description, randomization, statistical analysis, standard errors for *t*'s, missing data, and efficiency. Lists of references and detailed experimental plans appear at the end of each chapter. The frequent numerical examples in the text encourage penetrating analyses, far beyond the initial *F*'s of the analyses of variance. Little special comment is made on the necessarily a posteriori nature of some of these analyses. Confounding is systematically dealt with in chapter VI, and summary tables of available confounded designs are conveniently given along with indices of the locus and degree of reduction in efficiency due to confounding. Chapters VII–XIII deal with confounding and balanced designs: split-plots, quasi-latin squares, balanced and partially balanced incomplete blocks, lattices, lattice squares, and incomplete latin squares. These chapters (pp. 218–390) contain numerous examples, details of numerous plans or layouts, and indexes to these plans. The material is presented with little repetition and with abundant cross-reference, so that the careful reader can almost use the text in cookbook fashion.

In chapter XIV, the analysis of a series of experiments is attempted. The steps here advocated require experience and "nerve." Procedures to follow when the standard assumptions fail are outlined briefly, but the rationale is consistently omitted. The last chapter gives tables of random permutations of 9 and 16 numbers with instructions on their use, and the book concludes with a selected bibliography and the usual index.

Psychologists should profit immensely from this fine book. Long before they heard of Fisher, psychological researchers used orthogonality and balancing procedures in their experiments. Recently they have learned something about the proper statistical analysis of their

designs. Another less intuitively obvious set of procedures is available in confounding and balancing, and Cochran and Cox make these more available to the psychologist than do any previous authors. The reviewer predicts that extensive use will be made of the balanced designs—chiefly of the squares, latin, lattice, Youden, etc. These designs will be especially useful when a multiplicity of treatments cannot all be given to the same experimental subject because of time, fatigue, transfer effects, etc. *Experimental Designs* opens the door to what has been a relatively inaccessible area to most of us, and the authors merit our thanks. It is the reviewer's impression that the more the reader understands the basic principles of design, the more he will applaud the achievement represented by this book.

DAVID A. GRANT.

University of Wisconsin.

HARRIS, J. D. *Some relations between vision and audition.* Springfield, Ill.: Charles C Thomas, 1950. Pp. v+56. $1.50.

When sense modalities are being compared, relations of difference are more easily established than those of similarity. But many sensory phenomena transcend modal boundaries, and some, such as adaptation, appear in all sense departments. This little book, consisting of a dozen or so brief essays (the majority less than a thousand words in length), stresses the functional resemblances between vision and audition as a start "toward arriving at general principles of organization and theory of the whole sensorium."

The points of comparison of the two senses, not entirely systematically selected, themselves exemplify the difficulties inherent in making intermodal comparisons. The topics are: absolute and differential sensitivity, the relative ranges of intensities mediated, wave frequency as a sensation determinant, relative efficiencies as energy integrators, ways in which sensations develop and decay, responses to regularly interrupted stimuli (flicker vs. flutter), the question of bilateral interaction, events in single nerve fibers, central and peripheral determinants of acuity, quantum considerations, and intersensory phenomena.

The experimental evidences cited are quite commodiously documented (a bibliography of over 100 titles). Still, there are a few topics in connection with which one could wish that the visual literature had been appealed to in as competent a manner as was the auditory.

FRANK A. GELDARD.

University of Virginia.

PASTORE, NICHOLAS. *The nature-nurture controversy.* New York: King's Crown Press, Columbia Univ., 1949. Pp. xvi+213. $3.25.

Pastore strongly believes that conservatives are hereditarians and that liberals are environmentalists. To support this thesis he examines

the writings of distinguished hereditarians and environmentalists and finds eleven out of twelve hereditarians to be political conservatives, and eleven out of twelve environmentalists to be political liberals. Pastore feels that the political attitudes determine the scientific views. To quote (p. 181): "Different scientists are differently affected by their political allegiances. . . . The nature-nurture controversy, qua controversy, has been sustained by the conflicting social philosophies of the scientists. . . . It is probable, however, that in most cases the scientists were not aware of the specific impact of their political loyalties upon their scientific thinking."

Eight specific comments can be made:

1. The selection of cases was made by Pastore, who admits that the sample is not statistically valid.

2. The other regression equation might well have been used by selecting liberals and conservatives and seeking their views on heredity and environment.

3. Since no additional judges were used, the results present Pastore's image of the views studied, not a consensus. He says (p. 17) that the cited quotations are intended to typify a particular point of view, but adds later that they are not meant to be representative of the total expression of the person quoted.

4. The effect of change with chronological age is neglected. Is the Watson of *Comparative Psychology* the same as the Watson of *Psychological Care of the Infant and Child*?

5. The era or social climate in which the person wrote is not controlled. In their own time, some whom Pastore labels as conservatives may have been liberal. All those studied lived after 1822, in the period of the greatest social change and the greatest advance in science. The hereditarians were all born prior to 1888, whereas four of the twelve environmentalists were born later than 1890. In Pastore's own classification of the periods in which the persons studied were influential (pp. 17–18), ten of the hereditarians start with 1900 as compared with four of the environmentalists, while two of the hereditarians start with 1919 as compared with eight for the environmentalists. What Pastore may be studying is the progressive modification of thinking as the complexity of the problem of heredity and environment has become more apparent with time.

6. In order to equate the data upon which the opinions studied are based, more control should be exercised over the intellectual and scholarly background of the persons studied. According to Pastore's classification, the hereditarians consist of six psychologists, three geneticists, two biologists, and one statistician, whereas the environmentalists consist of four psychologists, one anthropologist, three geneticists, one biologist, two sociologists, and one educator. This suggests that the type of data with which the person has had experience may determine

attitudes. Studies of nature-nurture may possibly determine political and social attitudes.

7. The objective effects of the work of the particular persons upon society is not discussed. Some who are listed as hereditarians have had a much greater impact in creating better atmospheres and good lives for human beings than some who are listed as environmentalists.

8. No attention is given to intermediate positions in either variable studied; a scientist must be classed either as an hereditarian or an environmentalist, as a liberal or a conservative. Many of the statements quoted, however, indicate intermediate positions. This is particularly true of the geneticists, whether or not they are placed in the conservative or liberal group.

The semantic problem with which Pastore struggles is a difficult one. But his solution is too easy. If his essential thesis is correct, namely, that a person's scientific views are a product of his political attitudes, in justice to his readers Pastore should state his own political attitudes, since they may be determining his own analysis of the data. From the book as a whole, one gets the feeling that Pastore believes that hereditarianism, conservatism, and badness are identical, while environmentalism, liberalism, and goodness are identical. This reminds one of the Stendler finding that first-grade children confuse beautiful, rich, and good and that as they mature, the concepts are differentiated.

JOHN E. ANDERSON.

Institute of Child Welfare, University of Minnesota.

HURLOCK, ELIZABETH B. *Child development.* (2nd Ed.) New York: McGraw-Hill, 1950. Pp. xvi+669. $4.50.

This book represents an extensive revision of a college text first published in 1942. The most significant change has been the abandonment of the longitudinal approach of the earlier work, in keeping with what the author states is "the recent trend in many colleges and universities to separate the study of child development from that of adolescence and to present the two in different courses" (p. ix). Because of the increasingly large amounts of research in both fields, she continues, it is impossible to cover the whole span of development in a single book. Therefore, all material pertaining to adolescence has been removed from this edition, and three new chapters have been added on childhood interests, family relationships, and personality.

The first two chapters present, respectively, a brief sketch of the history of child psycholgy and of the principles of development. Including the new material, the remaining thirteen chapters deal with the usual topics of child psychology, such as prenatal development, the newborn, physical growth, motor development, speech, play, the development of understanding, and emotional, social, and moral development. A bibliography is appended, comprised of some 1,250 items distributed over the last thirty years. In the relatively few instances

where explicit references are not given, statements are made which presumably reflect the general literature or the author's own views.

While the order of the chapters is fairly typical of texts of this sort, the basis for the arrangement of material within a chapter is not as easy to grasp. Furthermore, the inclusion of some systematic generalizations with reference to the major aspects of human behavior discussed and to their *development* would have added to the generally excellent presentation of the enumerated studies and would possibly have made it easier for the student to synthesize the citation of so many specific pieces of research. However, the liberal use of pictures, tables, graphs, charts, and other illustrative materials makes this text not only interesting but also a very readable elementary handbook in its area.

In terms of age range and subject matter this book is even more, as Dennis observed (*Psychol. Bull.*, 1943, **40**, 536–537) of the earlier work, "a college textbook in child psychology." It is one, however, that is clearly written, and it does present the student with well-digested, documented, and illustrated summaries of research from a vast literature on a wide variety of topics in the field of child psychology.

S. O. ROBERTS.

Fisk University.

VON BONIN, GERHARDT. *Essay on the cerebral cortex.* Springfield, Ill.: Charles C Thomas, 1950. Pp. xiv+150. $3.75.

Written by an active and leading student of neuroanatomy, this essay on the cerebral cortex emphasizes structural differentiations and relations as determined by anatomical and histological techniques. Physiological methods, behavioral data, and psychological analysis are accorded only nominal mention. The essay is divided into seven chapters: "Historical Introduction" (12 pp.), "Theoretical and Phylogenetic Introduction" (15 pp.), "General Morphology of the Cortex" (33 pp.), "Sensation" (33 pp.), "Action" (18 pp.), "Prediction" (7 pp.), "Emotion" (8 pp.). There are a few original suggestions—notably in regard to the extrapyramidal system and the somesthetic pathways—but in general the author maintains a conservative position and a modest, nondogmatic attitude.

From the author's basic view that the output of the brain "is a definable function of its input," it follows that sensation and action "are obvious chapter headings." But this concept, he points out, is inadequate in reference to the major portion of cortical activity, since man developed languages "based on a 'symbolific' function . . . which cannot be considered an output in the sense in which a simple muscular movement . . . is called an 'output.'" The central difficulty, of course, lies in the fact that psychological analysis, which "has to proceed independently of anatomy and physiology," has made little progress in de-

fining the factors or functions which neurology is to localize. Furthermore, "we have no right to expect that all the factors will be localizable in the sense in which classical tradition used that term." However, since "both the koniocortex of the primary sensory areas and the agranular cortex of the precentral motor cortex must remain unencumbered by things past in order to function efficiently," memory and prediction are localized, by elimination, in the homotypical cortex. Looking forward as well as looking backward "must be based on the possibility of interchanging spatial and temporal relations."

In respect to general principles of neurological investigation, von Bonin affirms that "the function of the cerebral cortex cannot be explained by one principle. . . . Two systems are of paramount importance, the 'random' intracortical and the 'ordered' corticocortical system. . . . Within the (macroscopic) corticocortical system the causal laws of macroscopic physical events, as we know them from mechanics, will hold good; within the (microscopic) intracortical system the laws will be statistical as in thermodynamics or like those which the physicists develop for 'fields,' as in electromagnetism."

The virtues of this little volume are a forceful demonstration of the driving force behind the "conviction that structure should be understood as an enduring order impressed upon a flow of energy." Its deficiencies perhaps reflect more the magnitude and inherent difficulties of the problem than any specific inadequacies of the present author.

HENRY W. NISSEN.

Yale University.

PARTEN, MILDRED. *Surveys, polls and samples: Practical procedures.* New York: Harper, 1950. Pp. 624. $5.00.

This book represents a comprehensive treatment of a field which, in spite of a considerable body of work, is still more art than science. On the whole, the author succeeds admirably in giving a balanced picture of the survey field, with emphasis on "practical procedures." The chapters on planning, methods of securing information, construction of the schedule or questionnaire, interview and mail questionnaire procedures, sources of bias and processing (editing, coding, and tabulation) provide an excellent (and practical) guide to these phases of survey operations. They should prove useful both as a text for courses in survey procedures and as a "check list" for survey directors looking for weak spots in their own operations and suggestions for modifications. The value of the book is considerably enhanced by an extensive bibliography.

On the debit side are the unsatisfactory treatment of sampling and the unqualified presentation as "fact" of many items which are, at best, "opinion." An example of the latter defect is the statement on page 161, "processing the data from booklets is awkward and time-consuming. For efficient processing it is almost essential that the information be

transcribed to a form better adapted to office procedures." No evidence for this conclusion is presented, and the reviewer knows of no conclusive demonstration that processing a well-designed booklet questionnaire is less efficient than processing any other type of questionnaire which contains the same amount of information.

Many of the assertions in the book are undoubtedly true. Most of them have not as yet been experimentally verified. Many readers will fail to grasp what is, I am sure, clear to the author of this book: there is considerable experience and extensive "folklore" in the survey field but only a meager amount of scientifically verified fact. The book has done a painstaking and admirable job of collecting a wealth of experience from the individuals and organizations in the survey field. The reader must continually bear in mind that he is reading a summary of accumulated experience and that uncritical application of the experience of others may be dangerous.

The treatment of sampling is quite disappointing. If the book is used as a textbook in a course on surveys, the instructor will do well to omit these chapters and draw on other sources. There is some excellent material presented (particularly in the chapter "Procedures for Drawing Samples"). On page 64, there is set forth an excellent statement of the aims of good sampling and the reviewer heartily endorses the statement (p. 403) that "research in the social sciences is still at a stage where biased errors are likely to be twenty to a hundred times as great as random errors unless the survey and sampling are very carefully designed and executed." In spite of these good features the sampling discussion is both confused and inaccurate.

Probably the most serious defect of the sampling discussion is the presentation of material based on the theory of unrestricted random sampling with the implication that it applies (at least reasonably well) to almost any type of sampling. While there are some cautions on this point, the general effect is to reinforce the careless procedures fostered by the inadequacies of most of the statistical textbooks used by psychologists. The discussion of sampling also perpetuates the fallacious identification of quota sampling with stratified sampling.

One point which will be forcibly driven home by a critical reading of this book is the inadequacy of published data about survey costs. The author, quite correctly, stresses the importance of cost considerations in planning surveys. While the discussion of survey costs does not explicitly indicate the paucity of cost data, the information presented speaks for itself. In passing, the reviewer questions the accuracy of the statement that "The Bureau of the Census has found payment per schedule the most satisfactory because enumerators have to make their calls when there is a chance of finding people at home."

There are other minor inaccuracies (and a few major ones) in this volume. These defects may be unavoidable in a book of this scope. In

spite of the defects, there is a wealth of valuable advice for the research worker who reads the book carefully *and* critically.

ELI S. MARKS.

U. S. Bureau of the Census.

FREEMAN, FRANK S. *Theory and practice of psychological testing.* New York: Holt, 1950. Pp. xxi+518. $3.50.

The purpose of this book is stated clearly in the preface: "In the developing scientific field of psychological testing, there is recurrent need for periodic presentation of comprehensive descriptions of tests. This volume is intended to meet such a need." The major portion of the book is, accordingly, devoted to descriptions of a large number of tests. Considerable space (54 pp.) is devoted to the Binet scales and their many revisions. Other chapters are devoted to "Semi-Performance and Performance Scales," "Nonverbal Group Scales of Mental Ability," "Verbal and Mixed Group Scales of Mental Ability," "Scales for Infants and Preschool Children," "Aptitude Tests," "Tests of Eductional Achievement," "Personality Inventories and Rating Scales," "Projective Methods," and "Situational Tests." In addition to the enumerations and descriptions of tests provided in the chapters listed above, the book includes chapters on "Basic Principles: Theoretical and Clinical," "Statistics of Psychological Testing," "Definitions and Analyses of Intelligence," "Intelligence Tests as Clinical Instruments," and "Applications and Problems." The last is a sort of catch-all for brief presentation of topics like the nature-nurture problem and racial and national differences.

The book is well written and should be a very useful textbook for courses in educational and psychological testing. The material included seemed to the reviewer to be satisfactorily accurate, although the tone or emphasis sometimes struck him as inappropriate. However, this is to be expected as a result of differences in points of view and is not a criticism of the book. The author is considerably more lenient with respect to the tests he discusses than the reviewer would be. It is entirely possible that in an elementary text materials that deserve severe criticism should simply be ignored; otherwise, students may become confused. The reviewer would like to have seen somewhat greater emphasis given to the flaws and limitations of tests such as the Bellevue Scales and the Chicago Tests of Primary Mental Abilities. As it is, the misinformation and speculation in Wechsler's *Measurement of Adult Intelligence* is never called to the student's attention. And the spuriously high reliability coefficients reported in the manual for the Chicago Tests of Primary Mental Abilities are accepted at face value. Unfortunately, a competing set of tests, the Differential Aptitude Tests (published by the Psychological Corporation), for which parallel-forms reliability coefficients are provided, is not mentioned in the book.

It is gratifying to note that the author is cautious in accepting the claims made for tests based on the results of factor analyses. Although the reviewer is firmly convinced of the value of factorial techniques, he is worried by the uses to which these are sometimes put.

A couple of what seem to the reviewer to be non sequiturs were noticed. Both are of minor importance.

1. In the footnote on page 314, we find the statement: "But so far as AQ's above 100 are concerned, they are illogical and psychologically unsound, for actually no one can achieve above his capacity." It is no doubt desirable to point out the dangers of the AQ, but it seems a bit strong to indicate that all AQ's above 100 are illogical and psychologically unsound. Though it may be impossible, by definition, for anyone to achieve above his ultimate capacity in a given field, this does not preclude our obtaining meaningful AQ's above 100. For example, a child may show as much capacity on a certain mental test as the average child does at age 10–0 but score as high on a test of general science as the average child does at age 12–0 because of intense interest in the field and favorable opportunities to follow his interest. An AQ above 100 for this child would be quite meaningful.

2. On page 173, the author mentions a correlation coefficient between scores derived from the Pintner-Paterson Scale and the Stanford-Binet Scale in a group of gifted children, and concludes: "Furthermore, the very low coefficient of .23 suggests, also, that the performance scale is particularly inadequate for differentiating between performance levels of gifted children." To the reviewer the coefficient suggests that the distribution of scores for each test was probably greatly restricted. In any event, it would be possible to have two tests yielding uncorrelated scores that would each make excellent discrimination among gifted children.

FREDERICK B. DAVIS.

Hunter College.

HEPNER, HARRY W. *Psychology applied to life and work.* (2nd Ed.) New York: Prentice-Hall, 1950. Pp. viii+724. $4.50.

In this thorough revision of Hepner's book, more material has been added, the structure has been tightened, and new chapters are included on such topics as interviewing, counseling, supervision, family problems, and group dynamics.

The section on industrial psychology represents perhaps the most outstanding part of the book. Although the individual worker's role in the industrial environment is thoroughly covered, emphasis is placed on the implications of the important group research findings, such as the Hawthorne investigations, the "group atmosphere" studies of Lewin, Lippitt, and White, the new reports from the University of Michigan's Survey Research Center, Leighton's work on morale, and the advances in sociometry and attitude measurement.

Most of Professor Hepner's many illustrations and examples are properly documented, but there are instances when a more critical

evaluation of the illustrations would have been desirable. General Motors' "Why I like my job contest," for instance, is described without a critique of some of its sampling procedures and assumptions (pp. 471–473). A statement that "union members are more inclined to favor 'holding back' (of production) than are non-union workers" (p. 401) needs support in view of the contradictory evidence recently provided by Selekman.

Similarly, the section on the adjustment concept contains some conclusions based on clinical insight which could have been more profitably supported by experimental evidence. The statement "to understand an adult we must understand *how he felt* as a child" (p. 113) requires examination in the light of findings of the field theoreticians. The statement that "the strength of a nation . . . depends mainly upon the strength of personality of its people" (p. 45) introduces a vaguely mystical conceptualization. The belief that "the human race has progressed because heroic individuals have used positive adjustment in the face of adverse situations" (p. 44) is not particularly strengthened by the case history of a life-term prisoner who was able to make good in the advertising business.

The above should not detract from the over-all excellence of the book. In his revision, Professor Hepner has improved an already good book, which no doubt will again enjoy wide popularity among teachers and students alike.

IRVING R. WESCHLER.

University of California, Los Angeles.

FROMM-REICHMANN, FRIEDA. *Principles of intensive psychotherapy.* Chicago: University of Chicago Press, 1950. Pp. xv+246. $3.75.

This book has many distinctive and admirable qualities. It is written with an almost epigrammatic conciseness and clarity. It is calm and objective in mood yet definite, confident, and provocative. Indeed, it is a daring book in that it describes many technical innovations and serenely disregards entrenched dogma and doctrine.

Throughout the book one senses the long and varied experience, the wisdom, and the expertness of the author. Useful observations are made regarding almost every type of personality deviation which a psychotherapist is likely to encounter; but the author is particularly interested in the intensive treatment of psychotics, and it is here that some of her most trenchant and original discussions are to be found. Outstanding in this connection is the author's demonstrated premise that psychotics are treatable on the basis of the same underlying principles as are applicable in the neuroses and that the "security operations," or symptoms, of these two classes of patients are not basically different. To many readers the discussion of the qualifications and vicissitudes of the therapist as a person will also be of special interest and value.

The author reflects, and has importantly contributed to, the thinking of the late Dr. Harry Stack Sullivan and other psychiatrists, psychologists, and social scientists who have been connected with Chestnut Lodge Sanitarium, the Washington School of Psychiatry, and the William Alanson White Institute of Psychiatry in New York. In keeping with the general position of the "Washington School," Dr. Fromm-Reichmann accepts many of the basic tenets of Sigmund Freud: the reality of repression (Sullivan preferred the term "dissociation"), the anxiety theory of symptom formation, the therapeutic significance and utility of transference ("parataxic distortion" of interpersonal relations), and the premise that the primal pathogenic conflict is between cultural pressures for conformity to moral standards and the biologically given impulses of sex and aggression. On the other hand, Dr. Fromm-Reichmann, like many of her colleagues, swiftly dismisses the speculations of Freud regarding the "death instinct" and libido theory, expresses serious reservations regarding the classical Oedipus complex, and breaks with established traditions regarding many of the more routine features of psychoanalytic procedure.

This book is so simply written that beginners in psychotherapy will find it a useful and inspiring guide, and it is so replete with shrewd insights and general clinical sophistication that few experts will fail to enjoy and profit from it. From the standpoint of the reviewer, the book has only two major defects. In what we may suppose was the author's wish to give a semantically streamlined and highly operational account of psychotherapy, the book at times descends to a level which is slightly reminiscent of cookbook rules. This shortcoming will be readily apparent. The other one is more difficult to identify. It will probably be first perceived as a lack of theory and conceptual integration. Many of the author's observations are laid end to end without being tied together in any systematic way. But as one thinks further about the book, it will become evident that this weakness is related, somewhat adventitiously, to the book's greatest strength. As already indicated, the author follows the lead of Sullivan in discarding most of the biological baggage that psychiatry and even psychoanalysis (in a little different guise) has traditionally carried. Dr. Fromm-Reichmann places the central problems of psychotherapy squarely in the context of interpersonal relations, where they assuredly belong. However, the author's case would have been materially strengthened if she had not also disregarded certain of the constructs, or intervening variables, which form so important a part of Freud's later theorizing. From a formally theoretical standpoint, the book as it stands represents a somewhat naive form of social behaviorism; but this is a defect which could be easily remedied and is more than compensated by the book's many virtues and strengths.

O. H. MOWRER.

University of Illinois.

LEFEVER, D. WELTY, TURRELL, ARCHIE M., & WEITZEL, HENRY I. *Principles and techniques of guidance.* (Rev. Ed.) New York: Ronald, 1950. Pp. vii+577. $4.25.

This ambitious volume of fifteen chapters is offered as a thoroughgoing revision of the first edition with the addition of two new approaches. First, while subscribing to the role of the guidance specialist, more attention is given the part played by the classroom teacher in the guidance program. Second, vocational guidance receives increased emphasis in the revised volume. The fifteen chapters are arranged into four parts. Part I deals with the basic concepts and principles of guidance; Part II with the group approach to guidance; Part III with individual guidance and counseling; and Part IV with evaluation and follow-up.

Contents of the book vary considerably among the chapters both in quality and detail of treatment. There is appropriate emphasis on the need for personalization and individualization of the curriculum. The usual techniques and tools of guidance are described and illustrated, some more thoroughly than others. The discussion of basic principles seems to the reviewer to be fundamentally sound, and there is an especially good chapter on tests. One chapter on selected guidance systems is timely, and will be helpful to those struggling with problems of organization.

The need for the guidance program, according to the authors, stems from the absence of an ideal system of education, and hence is recognized as a temporary expedient. Group approaches to guidance presently are therefore necessary. One chief function of the counselor is that of directing and supplementing the work of the teacher. It is not clear to the reviewer whether the authors would subscribe to the need for the counseling specialist under an ideal system of education.

Brief reference is made to nondirective counseling. It is recommended that "autocratic, prescriptive, and directive approaches" should be replaced by "some of the techniques" of the nondirective method. Just what of the techniques could be used without modification in the highly authoritative guidance setting presented by the authors, is not indicated. Elsewhere in the book it is suggested that techniques and methods used should lie somewhere between more highly structured and nondirective methods, which is hardly more clarifying.

While the authors regard the part-time or full-time counseling specialist as playing the double role of leader of the guidance program and specialist in dealing directly with certain adjustment problems, there is an absence of suggested framework within which the counseling specialist could coordinate the guidance efforts of classroom teachers and other staff members in meeting without duplication of effort the needs of the individual student. This, in the reviewer's opinion, is the greatest weakness of the book.

The guidance needs of exceptional students are recognized, and some of them are dealt with in detail. This discussion, however, has

limited application to those of superior intelligence and special talents, to whom a total of only two and a half pages are devoted. Priority is given to students with various kinds of problems—recognition of a problem being necessary to set the guidance machinery in motion. The authors are not alone in failing to give sufficiently detailed attention to the needs of those who present no obvious problems, but whose superior capacities need direction into stimulating outlets. In spite of its limitations the book has much to offer the discriminating reader.

JOSEPH V. HANNA.

New York University.

KENT, GRACE H. *Mental tests in clinics for children.* New York: D. Van Nostrand, 1950. Pp. ix+180. $2.45.

The author states that this book is intended for advanced students in clinical psychology. She states further that it was her wish "to present a point of view not generally recognized." This reviewer thinks that the author has achieved both goals with unusual success. This book will be of considerable help and interest to advanced students in clinical psychology, and the point of view, while not generally recognized, needs to be emphasized.

In chapter I, Dr. Kent introduces the problem of mental tests for children by considering the reasons for children's being referred to a clinic. The reasons discussed include mental retardation, reading difficulties, adoption suitability, and misconduct. The clinical examination is discussed in chapter II. The author emphasizes the need to make the examination appeal to the child, the use of praise, and the clinical observations.

Chapter III, "Sources of Error in Use of Tests," might be considered as the author's statement of rationale. The point of view is that undue emphasis and importance have been given to the quantitative aspects of mental testing. Diagnostic errors may derive from taking scores too literally. This emphasis is the major contribution of this book. For example, in chapter IV, "The Composite Scale Versus the Battery," the author states, "There is naturally a demand for the composite scale, on the ground that it can be presented more or less mechanically by a psychometrist of comparatively little training. It calls for more judgment on the examiner's part to select sub-tests suitable for a given subject than it requires to run through a rigidly prescribed series of tasks. It must not be forgotten, however, that the results of the mechanically presented examination are frequently misleading."

This reviewer finds the emphasis on the need to base a diagnosis on the complete appraisal of the child, rather than primarily on his scores, to be refreshing and stimulating. Some psychologists will not agree with this point of view, but Dr. Kent speaks authoritatively, and her discussion reveals a wealth of clinical experience from which her arguments derive. It is a practical and timely book. Clinical psychologists in service clinics and those responsible for training will find it helpful.

HELMER R. MYKLEBUST.

Northwestern University.

HAHN, M. E., & MACLEAN, M. S. *General clinical counseling.* New York: McGraw-Hill, 1950. Pp. xi+375. $3.50.

Written as a text for graduate students preparing to become professional counselors, this book discusses the principles and functions of counseling and describes some of the basic counseling techniques. The authors' approach is problem-centered or, perhaps better, centered about students with problems, and the general role of counseling in education for personality development is given little attention. Accepting this limitation, *General Clinical Counseling* is a commendable text.

The first part of the book concerns itself with a definition of counselors and counseling. The problem orientation is obvious in this definition: "Clinical counseling, as the term is used in this book, is a process which takes place in a 1-to-1 relationship between an individual troubled by personal problems with which he has been unable to cope alone and a professional personnel worker whose training and experience qualify him to aid others to reach solutions to various types of personal difficulties." The 1-to-1 process is emphasized to the point that counseling involving two or more clients—for instance, marriage counseling—falls outside the definition of clinical counseling.

Included in this section is a discussion of the preparation desirable for counseling. Conspicuous through its absence is the statement issued by the National Vocational Guidance Association in 1949.

Next follow sections on principles of clinical counseling and the nature of various types of problems, including vocational problems and educational problems. Problems with broader implications are not included in this section nor discussed systematically in the book.

The major portion of the book is devoted to a discussion of the tools and techniques of counseling. These include counseling records, the interview, standardized tests, sociometric devices, occupational information, and statistics. It seemed to the reviewer that undue attention was given to the Minnesota Occupational Rating Scale. Although this scale has been found to be extremely useful in the training of counselors, its actual use in counseling interviews is relatively limited.

The final section of the book is devoted to a systematic case study. The case study would be of greater instructional value if it included actual excerpts from the interviews. Before interviews could be recorded and accurate typescripts obtained, a textual description of case work was the best that could be used for instructional purposes, but now that recording equipment is so easily available, samples of the actual interviews can be used most effectively.

The chief criticism that can be made of this text is that it looks backward rather than forward. Most of the discussion centers about what counseling has been, rather than what it could be. Much of the relevant research done in counseling, and certainly there has been relatively little, is omitted, and the authors apparently do not base their presentation on that research. Such pertinent research as that recently coming from the University of Chicago is not mentioned. The recent emphasis on the social context in which counseling takes place also fails to appear in this book, although the recent contributions of

semantics and sociometry are recognized. The social significance of counseling is not explored.

Worth particular mention is the book's fine and lucid style of expression. It provides both interesting and pleasurable reading.

RALPH F. BERDIE.

University of Minnesota.

LANDIS, CARNEY, & BOLLES, M. MARJORIE. *Textbook of abnormal psychology.* (2nd Ed.) New York: Macmillan, 1950. Pp. x+634. $5.00.

According to its authors, "The present revision was undertaken because of the vast amount of new research findings in the medical, social, and biological sciences during and since World War II." The writers have interwoven these recent findings into the fabric of classical abnormal psychology in such an excellent fashion that one cannot tell where the old leaves off and the new begins. As was true of the original edition, published just five years ago, this revision is certainly one of the best treatments of abnormal psychology available today. In some ways it is the best.

There appear to be only a few important studies and recent trends omitted in the revision, but in the reviewer's opinion certain such omissions are to be regretted. For example, in the discussion of typology the authors cite Kretschmer's classical types but do not mention the methodologically improved studies of Sheldon. The recent elaborate factorial studies of Cattell and Eysenck are not treated, even though factor analysis appears to be one of the more important methodological advancements in the field of abnormal psychology. The very recent cooperative endocrine studies of the Worcester group under the direction of Hoagland and Pincus are not mentioned, despite the fact that these studies offer one of the most promising attacks on all the psychoses. The significant aspects of role-taking as elaborated by Cameron are overlooked, as are the recent attempts to bring behavior disorders into the theoretical framework of learning theory. There are a few other omissions although in the reviewer's opinion they are not so significant as those mentioned above.

The order of presenting the material in the revision is much better and far more "teachable" than was the case in the original edition. The book is arranged into five basic sections. The first section presents a brief orientation and introduces the reader to the eclectic viewpoint which the authors attempt to maintain throughout their book. The second and longest section is given over to the precise description of the various representative kinds of abnormal behavior. The order of presenting the descriptions keeps the various headings such as neurosis and psychosis distinct from each other. In the first edition this was not the case, and jumping back and forth was often quite confusing. The third

section presents an extremely fine review of the explanations of abnormal behavior dealing with such varied topics as heredity, culture, development, the brain, etc. That these authors are clearly at home with their material is quite evident from an examination of this section. The fourth section deals with the general features of psychopathology, which are classically treated under the heading of disorders of sensation, emotion, intellect, etc. The fifth and last section, headed "Diagnosis and Therapy," is improperly named since it includes much more than is generally considered under these specific topics. The reviewer feels that too little attention is given to diagnosis and therapy, even though these are usually topics included in so-called "clinical psychology." The chapter on the law as applied to abnormal psychology is especially good and should be carefully studied by all psychologists.

Though the authors state clearly that they intended to approach abnormal psychology from an eclectic and unbiased viewpoint, there is an unmistakable leaning toward biological interpretation. This leaning, in the reviewer's opinion, is easily defended in view of the greater number of precise and objective biologically oriented studies of abnormal behavior than can be found in the social field. Nonetheless, the authors have maintained an exceptional degree of open-mindedness while still preserving the skepticism of the scientist.

The authors state that if they "be allowed a prediction, it is that the next five to ten years will provide more additions to our basic knowledge of mental abnormality than have been made during the past fifty years." It is the reviewer's conviction that should the authors' prediction turn out to be true, they could do no greater service to their profession than to revise their book again.

ROBERT P. FISCHER.

Marietta College.

WATKINS, J. G. *Hypnotherapy of war neuroses.* New York: Ronald, 1949. Pp. x+384. $5.00.

With incisive style, Watkins writes about the skillful use of hypnotherapy with diverse problems. The book is of definite value to those interested in hypnosis, either for research or more pragmatic purposes. Part I is somewhat repetitious of well-known facts about neurosis, but Parts II, III, and IV discuss in detail techniques of induction, palliative therapy with individual cases, and advanced therapy with an entrenched phobia.

Whereas the more nondirectively oriented reader will be shocked by some of the procedures the author employs, the latter is frank about the limitations of hypnotic and directive methods generally. His basic aim in the military setting is not "to rebuild the patient's basic personality structure but . . . to relax, if not resolve, a few of the major unconscious conflicts and to strengthen the ego . . ." This reviewer was partic-

ularly impressed by the skill with which various nuances of hypnotic behavior were employed. Among these may be cited the "in-and-out" method wherein the patient is successively brought into and out of trance, progressively recalled memories, automatic writing, and re-adjusted memories. (A critical point, however, is that "adjusted" memory may be ineffective in that it may not provide a completely affective working through. Although this limitation is acknowledged, more research is needed on the length of time the changed memory persists, how well it is integrated, etc., before the method becomes widely practiced as a technique.)

One may take issue with the author's contention that antisocial behavior may be induced under hypnosis, although he does admit the evidence is controversial. R. W. White, for one, seems to favor the view that such behavior cannot be suggested. This is certainly an area where research must be pursued to shed important light on the superego.

A minor criticism is the insertion of a glossary defining obvious terms, which is redundant since the book is geared for the professional. However, this does not detract from a readable and useful casebook.

RALPH D. NORMAN.

University of New Mexico.